Basic Concept of Biotechnology

Editors

Dr. K.N.Chandrashekara

Dr. Ashok Yakkaldevi

FIRST EDITION

LAXMI BOOK PUBLICATION

258/34, RaviwarPeth,

Solapur-413005

Cell: +91 9595359435

Rs: /-

"Basic Concept of Biotechnology"

Dr. K.N. Chandrashekara

Dr. Ashok Yakkaldevi

© 2015 by Laxmi Book Publication, Solapur

All rights reserved. No part of this book may be reproduced in any form, by mimeograph or any other means, without permission in writing from the publisher.

ISBN-

Published by,

Laxmi Book Publication,
258/34, RaviwarPeth,
Solapur, Maharashtra, India.

Contact No. : +91 9595 359 435

Website: http://www.isrj.org

Email ID: ayisrj@yahoo.in

PREFACE

Biotechnology is broadly defined in a 1991 Office of Technology Assessment report as "any technique that uses living organisms (or parts of organisms) to make or modify products, to improve plants or animals, or to develop microorganisms for specific uses." This technology has been instrumental in the development and implementation of processes for the manufacture of antibiotics and other pharmaceuticals, industrial sugars, alcohols, amino acids and other organic acids, foods, and specialty products through the application of microbiology, fermentation, enzymes, animal cell and separation technology. Engineers, working with life scientists, often achieved scale-up to industrial production in remarkably short periods. A relatively small number helped to catalyze, over a period of 50 years, the growth of the pharmaceutical, food, agricultural-processing, and specialty-product sectors of the Indian economy to the point where sales now exceed $500 billion/year.

The past decades have witnessed an enormous development in biotechnology with regard not only to the isolation, synthesis, structure identification, and elucidation of the mode of action of molecules, but also to their application as tools within the life sciences. Biomolecules have proved to be of interest not only in biochemistry, but also in chemistry, biology, pharmacology, medicinal chemistry, biotechnology, and gene technology.

We are aware however that, despite all our efforts, it is impossible to include all aspects of biotechnology research in one book. We are not under the illusion that the text, although carefully prepared, is completely free of errors. Indeed, some colleagues and readers might feel that the choice of priorities, the treatment of different aspects of biotechnology research, or the depth of presentation may not always be as expected. In any case, comments, criticisms and suggestions are appreciated and highly welcome for further editions.

The editors, authors and publisher are pleased to present the book on *Basic Concepts of Biotechnology*. After years of studying the

individual components of living systems, we can now study the systems themselves in comprehensive scope and in exquisite molecular detail. We therefore face the tasks of effectively employing new technologies, of dealing with mountains of data, and, most important, of adjusting our thinking to understand complex systems as opposed to their individual components.

Basic Concepts of Biotechnology had its origins from Dr. Ashok Yakkaldevi, Laxmi Book Publication, India. This is a book for beginners. My goal here is to familiarize the inexperienced reader with the important of tools and techniques of biochemistry and biotechnology. Thus the description of certain instrumentation and applications is not highly rigorous. This book is not intended to be a laboratory manual or a compilation of the latest techniques. There are several excellent volumes available that provide more detailed descriptions of protein analytical techniques, mass spectrometry instrumentation and techniques, and applications of these technologies. The evolution of methods and applications in this area is now so rapid that no book really could be truly up-to-date. What is exciting about my experience in introducing *Basic Concepts in Biotechnology* to colleagues has been the creativity with which they then apply these tools. Ultimately, the exciting potential of biotechnology rests with those who can put new technologies to work to address important questions.

I would like to thank all the contribution authors, reviewers and publisher, who provided valuable suggestions, read and commented on several drafts of book chapters. I thank Dr. Ashok Yakkaldevi for excellent secretarial assistance. Finally, I thank my wife Nalina and my daughter Hamsini for their patience with me.

Dr. K. N. Chandrashekara

ACKNOWLEDGMENTS

The Dr. K. N. Chandrashekara thank Dr. Ashok Yakkaldevi for his thoroughly professional approach. His constant interest and input have had a significant impact on the final structure of this book. We are very grateful to the Scientists who have read and commented on various parts of this book: Dr. K. Chandrashekar, Dr. J. Mohan, Dr. S. Ganeshmurthy, Dr. M. K. Prasannakumar, Dr. L. N. Reddy, Dr. G. Sujan and Mr. K. N. Jagadish. All were responsible for considerable improvement to the text.

We are indebted to Dr. B. Radhakrishnan, Director, UPASI TRF TRI and many colleagues who generously provided reprints of publications or information and clarified areas of confusion. Whose support and encouragement during the preparation of this book is greatly appreciated. We sincerely express our appreciation to all the authors who contributed to this book. We are thankful to all the reviewers who helped, read through and verified the existing information in this book. Without those tireless efforts and patient support this book would not have reached the readers.

I am very grateful to Dr. Ashok Yakkaldevi for their thoughtful and meticulous work in efficient typesetting and proofreading the text. Suggestions from students and colleagues have been most helpful in the compiling of this book. We look forward to receiving similar input in the future.

Last but not the least; we are thankful to Laxmi Book Publication for accepting our proposal to publish this book, without their efforts and support this book would not have reached the readers.

List of Authors

K. N. Chandrashekara
Senior Scientist and Head
Division of Plant Physiology & Biotechnology
UPASI Tea Research Foundation Tea Research Institute
Valparai, Coimbatore, Tamil Nadu, India
knchandu1@gmail.com

M. Nalina
Senir Research Fellow and Ph.D Scholar
Division of Plant Physiology & Biotechnology
UPASI Tea Research Foundation Tea Research Institute
Valparai, Coimbatore, Tamil Nadu, India
nalinam08@gmail.com

A. Raghavendra
Division of Insect Ecology
National Bureau of Agricultural Insect Resources
Bangalore, Karnataka, India
raghuckm9@gmail.com

M. Chakravarthi
Senior Research Fellow and Ph.D Scholar
Division of Crop Improvement
Sugarcane Breeding Institute (ICAR)
Coimbatore, Tamil Nadu, India
chakra3558@gmail.com

P. Harunipriya
Research Associate
Division of Crop Improvement
Sugarcane Breeding Institute
Coimbatore, Tamil Nadu, India
harunibct@gmail.com

C. Brindha
Ph.D Scholar
Division of Crop Improvement
Sugarcane Breeding Institute (ICAR)
Coimbatore, Tamil Nadu, India
brindha.b@gmail.com

S. Vinoth
Senior Research Fellow
Department of Plant Science
Bharathidasan University
Tiruchirappalli, Tamil Nadu, India
vinogenes@gmail.com

M.S. Suma
Senior Research Fellow
Department of Mechanical Engineering
Indian Institute of Science
Bangalore, Karnataka, India
suma.shivayogi@gmail.com

Shilpa R Raju
Senior Research Fellow and Ph.D Scholar
Department of Mechanical Engineering
Indian Institute of Science
Bangalore, Karnataka, India
shilpa171182@gmail.com

Madhuri Biradar
Teaching Assistant
Department of Applied Genetics
Karnatak University
Dharwad, Karnataka, India
madhuribiradar@gmail.com

Harendra Modak
Ph.D Scholar
Department of Applied Genetics
Karnatak University
Dharwad, Karnataka, India
harendra.ind@gmail.com

A. R. Chidanand
Research Associate and Ph.D Scholar
Department of Biotechnology
University of Agricultural Sciences Dharwad
Dharwad, Karnataka, India
chidanandiabt@gmail.com

M. Ranjith Kumar
Post Doctoral Researcher
Biotechnology Lab
Department of Horticulture

Prakash M Navale
Research Associate and Ph.D Scholar
Department of Biotechnology

Sunchon National University
Suncheon, South Korea
mrkumarbiotech@gmail.com

Indian Institute of Horticultural Research
Hessarghatta Lake Post
Bangalore, Karnataka, India
prakash.navale@rediffmail.com

H. D. Sowmya
Senior Research Fellow and Ph.D Scholar
Department of Biotechnology
Indian Institute of Horticultural Research
Hessarghatta Lake Post
Bangalore, Karnataka, India
sowmyaprakash97@gmail.com

Sandeep Telkar
Senior Research Fellow and Ph.D Scholar
Department of Biotechnology
Kuvempu University
Shimoga, Karnataka, India
sanbioinfo@gmail.com

Mr. S. Ashokraj
Ph.D Scholar
Biotechnology Lab
Department of Horticulture
Sunchon National University
Suncheon, South Korea
araj866@gmail.com

V. Brindha Priyadarisini
Assistant Professor
Department of Microbial Biotechnology
Bharathiar University
Coimbatore, Tamil Nadu, India
brindha02@rediffmail.com

Index

Chapter No.	Content	Author Name	Page No.
1.	Biomolecules	Nalina. M, Raghavendra. A and Chandrashekara. K.N	1-49
2.	Computer Applications and Biostatistics	Sandeep Telkar, Nalina. M, Sowmya. H.D, Prakash M Navale and Chandrashekara. K.N	50-84
3.	Macromolecules and Analytical Techniques	Raghavendra. A, Nalina. M and Chandrashekara. K.N	85-124
4.	Plant Molecular Farming: A Promising Stratergy In Biotechnology	Harunipriya. P, Chakravarthi. M, Brindha. C and Chandrashekara. K. N	125-182
5.	Plant Transgenics: Genetic Engineering Approch To Devlop Biotic Stress Resistance Plants	Chakravarthi. M, Vinoth. S and Chandrashekara. K.N	183-238
6.	Animal Biotechnology	Shilpa R Raju, Suma M.S, Nalina. M and Chandrashekara. K.N	239-321
7.	Medical Biotechnology	Harendra Modak, Madhuri Biradar,	322-345

		Nalina. M and Chandrashekara. K.N.	
8.	Tools and Techniques in Biotechnology	Harendra Modak, Chidanand Rabinal, Madhuri Biradar, Nalina. M and Chandrashekara. K.N.	346-416
9.	Antibiotics: Microbial sources, Production and Optimization	Ranjith Kumar. M, Ashokraj. S and Brindha Priyadarisini. V	417-457
10.	Biotechnology in Forensic Sciences	Harendra Modak and Madhuri Biradar	458-506

Chapter 1
Biomolecules

Nalina. M, Raghavendra. A and Chandrashekara K.N

A living system grows, sustains and reproduces itself. The most amazing thing about a living system is that it is composed of non-living atoms and molecules. The pursuit of knowledge of what goes on chemically within a living system falls in the domain of Biochemistry. Even though there are thousands of different types of molecules in a cell, there are only a few basic classes of bimolecular like carbohydrates, proteins, nucleic acids, lipids, etc. Proteins and carbohydrates are essential constituents of our food. In addition, some simple molecules like vitamins and mineral salts also play an important role in the functions of organisms. The complexity of even the simplest of life forms, the single cell, cannot be overstated. Nevertheless, from a chemical perspective, cellular components can be segregated into macromolecules (DNA, RNA, proteins, etc.), relatively simple molecules (amino acids, monosaccharide's, and lipids), and their precursors: CO_2, H_2O, and NH_3. In general, the macromolecules tend to be polymers of small bimolecular; however, each of these molecules, whether simple or complex, is involved in a myriad of intricate metabolic reactions. A case in point is the monosaccharide glucose which is synthesized from H_2O and CO_2. When degraded to its precursors, it provides the cell with its energy requirements for such diverse processes as macroscopic movement as well as the synthesis of complex macromolecules. In addition, glucose is the fundamental building block of macromolecules

such as starch and cellulose. This basic theme, in which the cell uses a simple small molecule in a multitude of processes, is typical of how relatively small bimolecules are used in living systems. In this chapter we will elaborate the chemistry, properties and metabolism of four bimolecules: amino acids, carbohydrates, lipids, and nucleotides and their roles in metabolism. You are aware of that our body, plants and other animals are made up of many chemical substances. There are certain complex organic molecules which form the basis of life. These build up living organisms and are also required for their growth and maintenance. Such molecules are called bimolecules. The main classes of bimolecules are carbohydrates, proteins, lipids, nucleic acids, enzymes, hormones etc. In this lesson, you will study about the structures and functions of some important bimolecules.

Carbohydrates

Carbohydrates are the most abundant bimolecule belonging to class of organic compounds found in living organisms on earth. Each year, more than 100 billion metric tons of CO2 and H2O are converted into cellulose and other plant products due to photosynthesis. Living matter is largely made of bimolecule consisting of water and complex polymers of amino acids, lipids, nucleotides and carbohydrates. Carbohydrates are most special of them in that they remain associated with the three other polymers mentioned. Carbohydrates are linked with amino acid polymers (proteins) forming glycoprotein's and with lipids as glycolipids. Carbohydrates are present in DNA and RNA, which are essentially polymers of D-ribose-phosphate and 2-deoxy-D-ribose phosphate to which purines and pyrimidines bases are attached at the C-1 reducing position. Carbohydrates are a widely diverse group of compounds that are ubiquitous in nature. More than 75% of the dry weight of the plant world is carbohydrate in nature - particularly

cellulose, hemicelluloses and lignin. Carbohydrates comprise a comprehensive group of naturally occurring substances, which include innumerable sugars and sugar derivatives, as well as high-molecular weight carbohydrates (polysaccharides) like starch and cellulose in plants and glycogen in animals. A polysaccharide molecule is composed of a large number of sugar or sugar-like units. Carbohydrates are of great importance in biology. The unique reaction, which makes life possible on Earth, namely the assimilation of the green plants, produces sugar, from which originate, not only all carbohydrates but, indirectly, also all other components of living organisms. Carbohydrates form a very large group of naturally occurring organic compounds which play a vital role in daily life. They are produced in plants by the process of photosynthesis. The most common carbohydrates are glucose, fructose, sucrose, starch, cellulose etc. Chemically, the carbohydrates may be defined as polyhydroxy aldehydes or ketones or substances which give such molecules on hydrolysis. Many carbohydrates are sweet in taste and all sweet carbohydrates are called as sugars. The chemical name of the most commonly used sugar in our home is sucrose.

Classification of Carbohydrates

Carbohydrates are classified into three groups depending upon their behavior on hydrolysis.

(i) Monosaccharide's:

A polyhydroxy aldehyde or ketone which cannot be hydrolyzed further to a smaller molecule containing these functional groups is known as a monosaccharide. About 20 monosaccharides occur in nature and glucose is the most common amongst them. Monosaccharides are further classified on the basis of the number of carbon atoms and the functional group present in them. If a monosaccharide contains an aldehyde group, it is known as an aldose and if it contains a keto group,

it is known as a ketose. The number of carbon atoms present is also included while classifying the compound as is evident from the examples given in Table 1.

Table 1: Classification of monosaccharides

No. of carbon atoms	Type of monosaccharide	
	Aldose	**Ketose**
Three	Aldotriose (Glyceraldehyde)	Ketotriose
Four	Aldotetrose ((Xylose)	Ketotetrose
Five	Aldopentose(Erythrose)	Ketopentose
Six	Aldohexose (Glucose)	Ketohexose
Seven	Aldoheptose	Ketoheptose

Glucose occurs freely in nature as well as in the combined form. It is present in sweet fruits and honey. Ripe grapes also contain glucose in large amounts.

(ii) Disaccharides:

Carbohydrates which give two monosaccharide molecules on hydrolysis are called disaccharides e.g. sucrose, maltose, lactose etc.

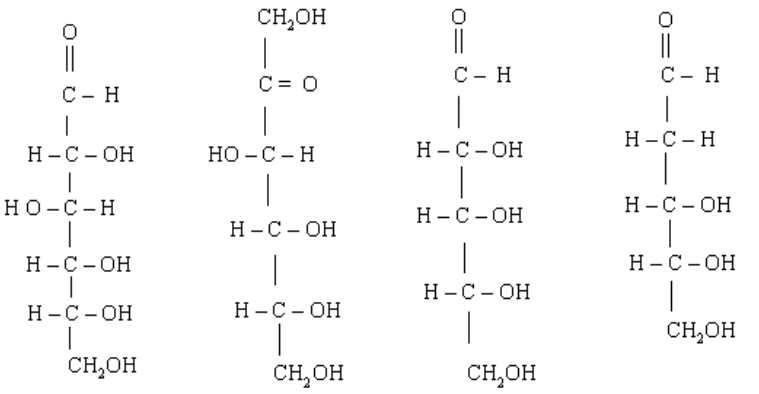
(iii) Polysaccharides:

Carbohydrates which yield a large number of monosaccharide units on hydrolysis e.g. starch, glycogen, cellulose etc

Structure of Monosaccharide's

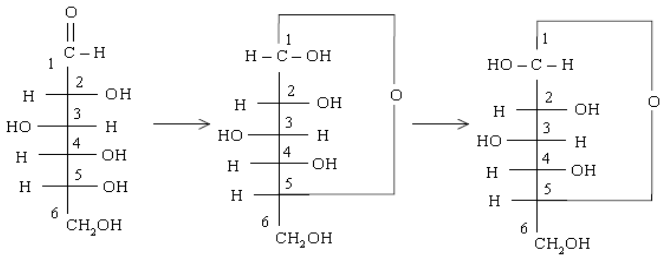
Although a large number of monosaccharide's are found in nature, we will confine our discussion here to four of them only viz. D-glucose, D-fructose, D-ribose and 2-deoxy-D-ribose. D-Glucose (an aldohexose) is the monomer for many other carbohydrates. Alone or in combination, glucose is probably the most abundant organic compound on the earth. D-Fructose (a ketohexose) is a sugar that is found with

glucose in honey and fruit juices. D-Ribose (an aldopentose) is found in ribonucleic acids (RNA) while. 2-Deoxy-D-ribose is an important



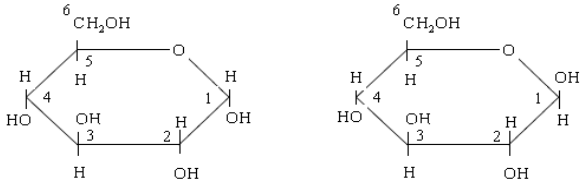
constituent of the deoxyribonucleic acids (DNA). Here, the prefix 2-Deoxy indicates that it lacks oxygen at carbon no. 2

These monosaccharides generally exist as cyclic compounds in nature. A ring is formed by a reaction between the carbonyl group and one of the hydroxyl groups present in the molecule. Glucose preferentially forms the six member rings which can be in two different isomeric forms called α- and ß-forms (shown below as I & II). The two forms differ only in the arrangement of the hydroxyl group at carbon No.1. Such isomers are called anomers. Formation of these cyclic structures (I and II) from the open chain structure can be shown as follows.



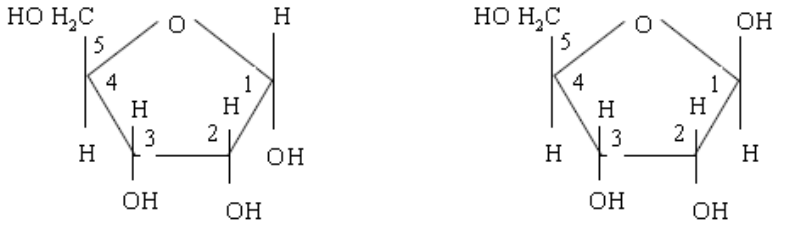
(I) α – D – glucose (II) β – D –glucose

The cyclic structures I and II are more appropriately represented as Ia and IIa.



(Ia) (IIa)

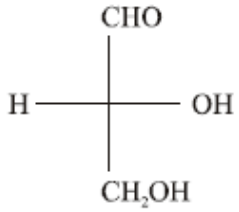
The α- and ß-forms of other sugars also exist in the cyclic form. D-Ribose forms a five member ring structure as shown below



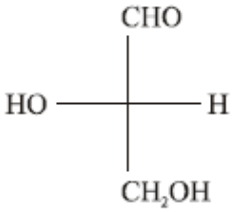
α – D – ribose β – D – ribose

D-before the name of above example indicates the configuration of particular stereoisomer. Stereoisomers are assigned relative configurations as D– or L –. This system of assigning the relative

configuration refers to their relation with glyceraldehydes. Glyceraldehydes contain one asymmetric carbon atom so exists in two enantiomeric forms as shown below.

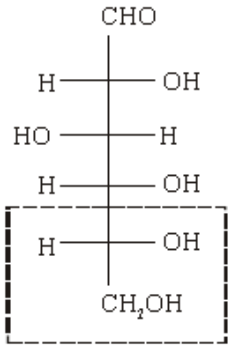


(+) –glycaraldehyde

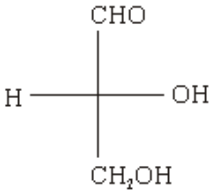


(–) –glyceraldehyde

All those compounds which can be correlated to (+) -glyceraldehyde are said to have D-configuration and those can be correlated to (–) -glyceraldehyde are said to have L–configuration. In monosaccharides it is the lowest asymmetric carbon atom (shown in the box) by which the correlation is made. As in (+) glucose the lowest asymmetric carbon atom has –OH group on the right side which matches with (+) glyceraldehyde hence it is assigned D-configuration.



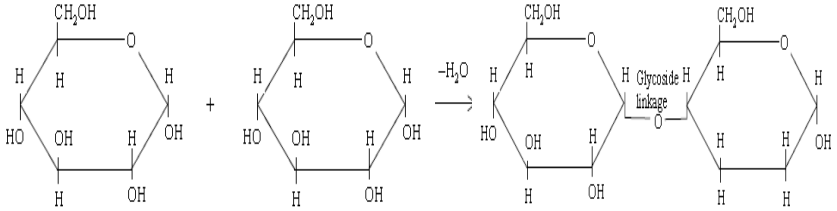
(+) – glucose or D-glucose



(+) – glyceraldehyde or D-glyceraldehyde

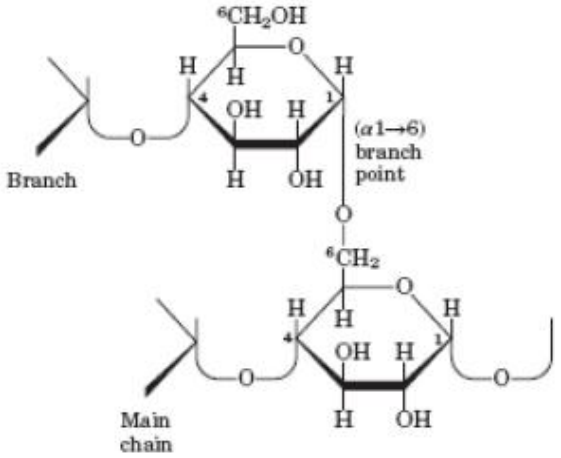
Structure of Di-Saccharides and Polysaccharides

Disaccharides are formed by the condensation of two monosaccharide molecules. These monosaccharides join together by the loss of a water molecule between one hydroxyl groups on each monosaccharide. Such a linkage, which joins the monosaccharide units together, is called glycoside linkage. If two α-glucose molecules are joined together, the disaccharide maltose is formed.

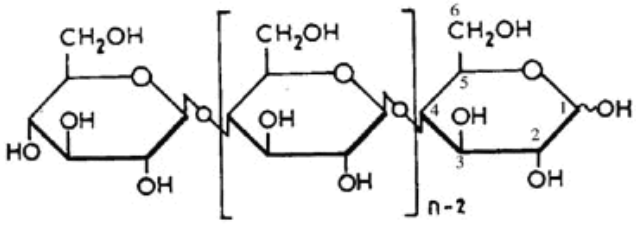


Similarly, sucrose (the common sugar) consists of one molecule of glucose and one molecule of fructose joined together. Lactose (or milk sugar) is found in milk and contains one molecule of glucose and one molecule of galactose. If a large number of monosaccharide units are joined together, we get polysaccharides. These are the most common carbohydrates found in nature. They have mainly one of the following two functions- either as food materials or as structural materials. Starch is the main food storage polysaccharide of plants. It is a polymer of α-glucose and consists of two types of chains- known as amylose and amylopectin. Amylose is a water soluble fraction of starch and is a linear polymer of α-D-glucose. On the other hand amylopectin is a water insoluble fraction and consists of branched chain of α-D-glucose.

The carbohydrates are stored in animal body as glycogen which is also a polymer of α -glucose and its structure is similar to amylopectin.



Cellulose is another natural polysaccharide which is the main component of wood and other plant materials. It consists of long chain of ß-D-glucose molecules.



Importance of carbohydrates

Carbohydrates are of great importance in biology. The unique reaction, which makes life possible on the Earth, namely the assimilation of the green plants, produces sugar, from which originate, not only all carbohydrates but, directly or indirectly, all other components of living

organisms. The carbohydrates are a major source of metabolic energy, both for plants and for animals that depend on plants for food. Aside from the sugars and starch that meet this vital nutritional role, carbohydrates also serve as a structural material (cellulose), a component of the energy transport compound ATP, recognition sites on cell surfaces, and one of three essential components of DNA and RNA. Importance can be considered under following headings;

Metabolic/Nutritional

The important role of carbohydrates, generally, in the metabolism of living organisms, is well known. The biological breakdown of carbohydrates (often spoken of as "combustion") supplies the principal part of the energy that every organism needs for various processes. Carbohydrates and their metabolism has been the subject of biochemical and medical research for a long time. Carbohydrates play a major role in promoting health fitness, form a major part of food and help a great deal in building body strength, by generating energy. They are one among the three prominent macronutrients that serve as excellent energy providers, the other two being fats and proteins. Carbohydrate intake can take place in different forms like sugar, starch, fibers etc., which are a dietary staple in most parts of the world, and the oxidation of carbohydrates is the central energy-yielding pathway in most non-photosynthetic organisms. The functions of carbohydrates are multiple and it is owing to this fact that it becomes all the more necessary to incorporate carbohydrates in our meal. For instant for energy generation, sugars and starch act as the perfect fuel that enables us to carry out our physical activities efficiently and effectively. Fiber does wonders in keeping your bowel function going smooth. Carbohydrates add on to the taste and appearance of food item, thus making the dish tempting and mouth watering. They are sometimes used

as flavors and sweeteners. Carbohydrates aid in regulating blood glucose and also do good to our body by breaking down fatty acids, thus preventing ketosis. Talking about the importance of carbohydrates, apart from its direct benefits, there is also an added advantage of carbohydrate consumption and that is that carbohydrates are found in different foods, which if eaten, also pave way for consuming other essential nutrients. Therefore, it is preferable to go in for distinctive carbohydrate food sources.

Biological importance

Ribose and 2-deoxyribose derivatives have an important role in biology. Among the most important derivatives are those with phosphate groups attached at the 5 position. Mono-, di-, and triphosphate forms are important, as well as 3-5 cyclic monophosphates. Purines and pyrimidines form an important class of compounds with ribose and deoxyribose. When these purine and pyrimidine derivatives are coupled to a ribose sugar, they are called nucleosides. In these compounds, the convention is to put a ′ (pronounced "prime") after the carbon numbers of the sugar, so that in nucleoside derivatives a name might include, for instance, the term "5′-monophosphate", meaning that the phosphate group is attached to the fifth carbon of the sugar, and not to the base. The bases are attached to the 1′ ribose carbon in the common nucleosides. Phosphorylated nucleosides are called nucleotides. One of the common bases is adenine (a purine derivative); coupled to ribose it is called adenosine; coupled to deoxyribose it is called deoxyadenosine. The 5′-triphosphate derivative of adenosine, commonly called ATP, for adenosine triphosphate, is an important energy transport molecule in cells. 2-Deoxyribose and ribose nucleotides are often found in unbranched 5′-3′ polymers. In these structures, the 3′carbon of one monomer unit is linked to a phosphate that is attached

to the 5′carbon of the next unit, and so on. These polymer chains often contain many millions of monomer units. Since long polymers have physical properties distinctly different from those of small molecules, they are called macromolecules. The sugar-phosphate-sugar chain is called the backbone of the polymer. One end of the backbone has a free 5′phosphate, and the other end has a free 3′OH group. The backbone structure is independent of which particular bases are attached to the individual sugars.

Genetic material in earthly life often contains poly 5′-3′, 2′-deoxyribose nucleotides, in structures called chromosomes, where each monomer is one of the nucleotides deoxy- adenine, thymine, guanine or cytosine. This material is commonly called deoxyribonucleic acid, or simply DNA for short. DNA in chromosomes forms very long helical structures containing two molecules with the backbones running in opposite directions on the outside of the helix and held together by hydrogen bonds between complementary nucleotide bases lying between the helical backbones. The lack of the 2′ hydroxyl group in DNA appears to allow the backbone the flexibility to assume the full conformation of the long double-helix, which involves not only the basic helix, but additional coiling necessary to fit these very long molecules into the very small volume of a cell nucleus. In contrast, very similar molecules, containing ribose instead of deoxyribose, and known generically as RNA, are known to form only relatively short double-helical complementary base paired structures. These are well known, for instance, in ribosomal RNA molecules and in transfer RNA (tRNA), where so-called hairpin structures from palindrome sequences within one molecule.

Rare sugars

Rare sugars are defined by the International Society of Rare Sugars (ISRS) as monosaccharide's and their derivatives that are rare in nature. They are hardly available for research purposes because of their expensiveness. "Izumoring", a structural framework containing all 34 six-carbon monosaccharide's linked by enzymatic reactions, has been proposed following the discovery of a key enzyme that converts abundantly occurring monosaccharide's in nature into rare sugars. This has made possible the mass production of rare sugars from inexpensive sugars such as D-glucose or D-fructose. Rare Sugars are mostly used in pharmaceuticals as precursors for a wide variety of carbohydrate-based drugs. These include nucleoside analogues, which are used in antiviral applications such as HIV, HBV and HCV. Another important class of compounds is complex oligosaccharides and olignonucleotides, which may be used as anti-inflammatory or anti-cancer agents, as well as in highly specific chronic pain relievers. They are also being used as precursors in the production of flavor chemicals, such as natural furan ones and Maillard reaction savory flavors. Furthermore, some of the rare sugars products have applications as nutraceuticals or they may be used in high-end cosmetic products. They are enlisted under the D & L series depending on their chirality.

Nucleic Acids

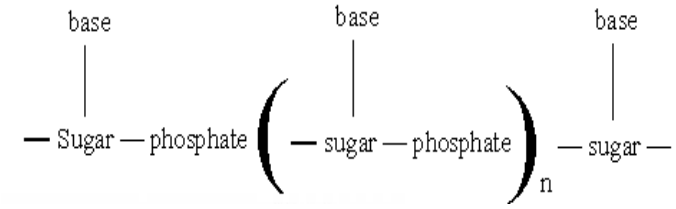
Every generation of each and every species resembles its ancestors in many ways. How are these characteristics transmitted from one generation to the next? It has been observed that nucleus of a living cell is responsible for this transmission of inherent characters, also called heredity. The particles in nucleus of the cell, responsible for heredity, are called chromosomes which are made up of proteins and another type of bimolecular called nucleic acids. These are mainly of two types, the

deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Since nucleic acids are long chain polymers of nucleotides, so they are also called polynucleotides. Why is a dog a dog and not a cat? Why do some people have blue or brown eyes and not black? From a chemical standpoint, how does the body know what particular type of protein is to be synthesized? How is this information transmitted from one generation to the next? The study of the chemistry of heredity is one of the most fascinating fields of research today. It was recognized in the 19th century that the nucleus of a living cell contains particles responsible for heredity, which were called chromosomes. In more recent years, it has been discovered that chromosomes are composed of nucleic acids. These are named so because they come from the nucleus of the cell and are acidic in nature. Two types of nucleic acids exist which are called DNA and RNA. They differ in their chemical composition as well as in functions. Like amino acids which are the building blocks of proteins, nucleotides are the building blocks of the nucleic acids, RNA and DNA. Aside from these major biological roles, nucleotides are important players in energy metabolism, coenzymes, and intermediary metabolism. Nucleotides are composed of a nitrogenous base, a sugar, and a phosphoryl group. Removal of phosphoryl group results in a compound known as a nucleoside. There are two types of bases found in nucleotides, purines and pyrimidines. The sugars are either D-ribose or 2-deoxy-D-ribose.

Structure of Nucleic Acids

Like all natural molecules, nucleic acids are linear polymeric molecules. They are chain like polymers of thousands of nucleotide units, hence they are also called polynucleotides. A nucleotide consists of three subunits: nitrogen containing heterocyclic aromatic compound (called

base), a pentose sugar and a molecule of phosphoric acid. So a nucleic acid chain is represented as shown below.



In DNA molecules, the sugar moiety is 2 -deoxyribose, where in RNA molecules it is ribose. In DNA, four bases have been found. They are adenine (A), guanine (G), cytosine (C) and thymine (T). The first three of these bases are found in RNA also but the fourth is uracil (U). The sequence of different nucleotides in DNA is termed as its primary structure. Like proteins, they also have secondary structure. DNA is a double stranded helix. Two nucleic acid chains are wound about each other and held together by hydrogen bonds between pairs of bases. The hydrogen bonds are specific between pairs of bases that are guanine and cytosine form hydrogen bonds with each other, whereas adenine forms hydrogen bonds with thymine. The two stands are complementary to each other. The overall secondary structure resembles a flexible ladder (Fig. 1.1). This structure for DNA was proposed by James Watson and Francis Crick in 1953.

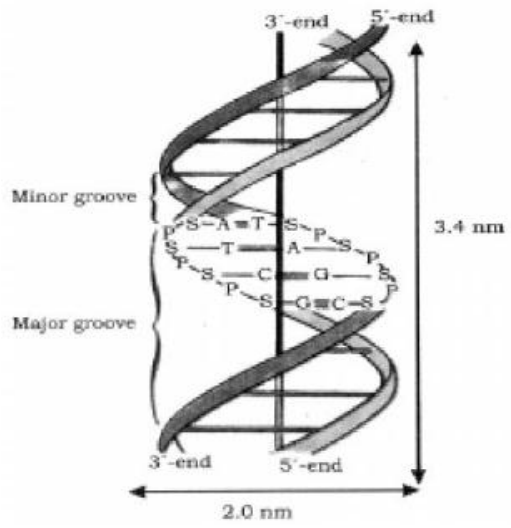


Fig.

1.1: Watson and Crick's double helix structure of DNA

Unlike DNA, RNA is a single stranded molecule, which may fold back on itself to form double helix structure by base pairing in a region where base sequences are complimentary. There are three types of RNA molecules which perform different functions. These are named as messenger RNA (m-RNA), ribosomal-RNA(r-RNA) and transfer RNA (t-RNA)

Biological Functions of Nucleic Acids

A DNA molecule is capable of self duplication during cell divisions. The process starts with the unwinding of the two chains in the parent DNA. As the two strands separate, each can serve as a master copy for the construction of a new partner. This is done by bringing the appropriate nucleotides in place and linking them together. Because the bases must be paired in a specific manner (adenine to thymine and guanine to cytosine), each newly built strand is not identical but complimentary to the old one. Thus when replication is completed, we have two DNA molecules, each identical to the original. Each of the new molecules is a double helix that has one old strand and one new strand to be transmitted to daughter cells (Fig. 1.2).

Another important function of nucleic acids is the protein synthesis. The specific sequence of bases in DNA represents coded information for the manufacture of specific proteins. In the process, the information from DNA is transmitted to another nucleic acid called messenger RNA, which leaves the nucleus and goes to the cytoplasm of the cell. Messenger RNA acts as template for the incorporation of amino acids in the proper sequence in protein. The amino acids are brought to the messenger RNA in the cell, by transfer RNA. Where they form peptide bonds. In short it can be said that DNA contains the coded message for protein synthesis whereas RNA actually carries out the synthesis of protein.

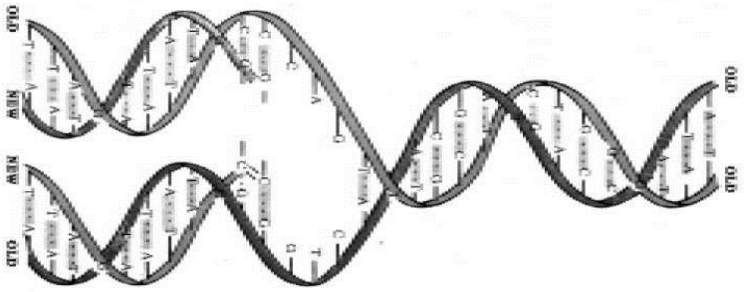


Fig. 1.2: Replication of DNA

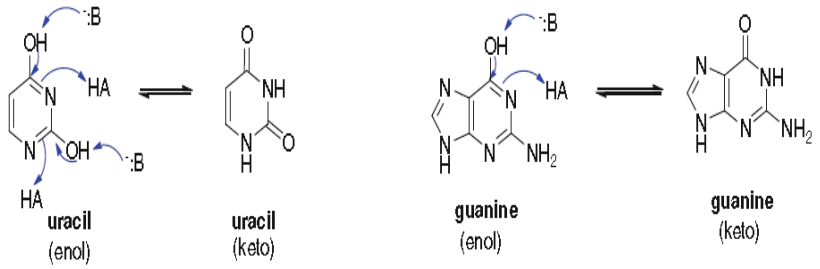
Nucleosides and nucleotides

Nucleosides are molecules formed by attaching a nucleobase to a ribose or deoxyribose ring. Examples of these include cytidine (C), uridine (U), adenosine (A), guanosine (G), thymidine (T) and inosine (I). Nucleosides can be phosphorylated by specific kinases in the cell, producing nucleotides. Both DNA and RNA are polymers, consisting of long, linear molecules assembled by polymerase enzymes from repeating structural units, or monomers, of mononucleotides. DNA uses the deoxynucleotides C, G, A, and T, while RNA uses the ribonucleotides (which have an extra hydroxyl (OH) group on the pentose ring) C, G, A, and U. Modified bases are fairly common (such as with methyl groups on the base ring), as found in ribosomal RNA or transfer RNAs or for discriminating the new from old strands of DNA after replication. Each nucleotide is made of an acyclic nitrogenous base, a pentose and one to three phosphate groups. They contain carbon, nitrogen, oxygen, hydrogen and phosphorus. They serve as sources of chemical energy (adenosine triphosphate and guanosine triphosphate), participate in cellular signaling (cyclic guanosine monophosphate and cyclic adenosine monophosphate), and are incorporated into important cofactors of

enzymatic reactions (coenzyme A, flavin adenine dinucleotide, flavin mononucleotide, and nicotinamide adenine dinucleotide phosphate).

Tautomerism

The nitrogenous bases in nucleosides and nucleotides, with the exception of adenine and adenosine, undergo enol-keto tautomerism. Studies have shown that the predominant species in solution is the keto form. Examples are uracil and guanine.



Biological Functions of Nucleic Acids

DNA is the chemical basis of heredity and may be regarded as the reserve of genetic information. DNA is exclusively responsible for maintaining the identity of different species of organisms over millions of years. A DNA molecule is capable of self duplication during cell division and identical DNA strands are transferred to daughter cells. Another important function of nucleic acids is the protein synthesis in the cell. Actually, the proteins are synthesised by various RNA molecules in the cell but the message for the synthesis of a particular protein is present in DNA.

Proteins

Proteins are the most abundant macromolecules in living cells. The name protein is derived from the Greek word 'proteios' meaning 'of prime importance'. These are high molecular mass complex amino acids. You will study about amino acids in the next section. Proteins are most essential class of biomolecules because they play the most important role in all biological processes. A living system contains thousands of different proteins for its various functions. In our every day food pulses, eggs, meat and milk are rich sources of proteins and are must for a balanced diet. Proteins are molecular tools that perform an astonishing variety of functions. In addition to serving as structural materials in all living organisms (e.g., actin and myosin in animal muscle cells), proteins are involved in such diverse functions as catalysis, metabolic regulation, transport, and defense. Proteins are composed of one or more polypeptides, unbranched polymers of 20 different amino acids. The genomes of most organisms specify the amino acid sequences of thousands or tens of thousands of proteins. Proteins are a diverse group of macromolecules. This diversity is directly related to the combinatorial possibilities of the 20 amino acid monomers. Amino acids can be theoretically linked to form protein molecules in any imaginable size or sequence.

An important reason for this remarkable discrepancy is demonstrated by the complex set of structural and functional properties of naturally occurring proteins that have evolved over billions of years in response to selection pressure. Among these are (1) structural features that make protein folding a relatively rapid and successful process, (2) the presence of binding sites that are specific for one or a small group of molecules, (3) an appropriate balance of structural flexibility and rigidity so that function is maintained, (4) surface structure that is appropriate for a protein's immediate environment (i.e., hydrophobic in membranes

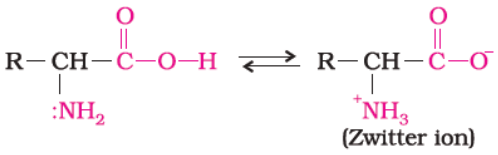
and hydrophilic in cytoplasm), and (5) vulnerability of proteins to degradation reactions when they become damaged or no longer useful.
Proteins can be distinguished based on their number of amino acids (called amino acid residues), their overall amino acyl composition, and their amino acid sequence. Molecules with molecular weights ranging from several thousand to several million daltons are called polypeptides. Those with low molecular weights, typically consisting of fewer than 50 amino acids, are called peptides. The term protein describes molecules with more than 50 amino acids. Each protein consists of one or more polypeptide chains.

Amino Acids

The hydrolysis of each polypeptide yields a set of amino acids, referred to as the molecule's amino acid composition. The structures of the 20 amino acids that are commonly found in naturally occurring polypeptides. Amino acids are the most versatile small biomolecules. They fulfil a number of extremely important roles in biology. These include: building blocks of proteins which are polymers of amino acids, precursors of hormones, and precursors of molecules with specialized physiological functions, e.g., the neurotransmitter dopamine and the hormone thyroxine are both derivatives of the amino acid tyrosine. As the name implies, amino acids contain amino and carboxyl groups. They can be divided into groups based on acidic, basic, and neutral properties when dissolved in water. They are also classified according to solubility, e.g., hydrophilic and hydrophobic. There are 20 so-called amino acids in proteins; however, one of these, proline, is in fact an imino acid. Nineteen of the 20 amino acids are optically active, i.e., they are capable of rotating plane polarized light either to the right (dextrorotary) or left (levorotary).

Classification of Amino Acids

Amino acids are classified as acidic, basic or neutral depending upon the relative number of amino and carboxyl groups in their molecule. Equal number of amino and carboxyl groups makes it neutral; more number of amino than carboxyl groups makes it basic and more carboxyl groups as compared to amino groups makes it acidic. The amino acids, which can be synthesized in the body, are known as nonessential amino acids. On the other hand, those which cannot be synthesized in the body and must be obtained through diet are known as essential amino acids (marked with asterisk in Table 14.2). Amino acids are usually colorless, crystalline solids. These are water-soluble, high melting solids and behave like salts rather than simple amines or carboxylic acids. This behavior is due to the presence of both acidic (carboxyl group) and basic (amino group) groups in the same molecule. In aqueous solution, the carboxyl group can lose a proton and amino group can accept a proton, giving rise to a dipolar ion known as twitter ion. This is neutral but contains both positive and negative charges.

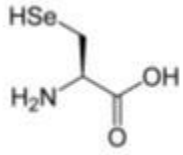


(Zwitter ion)

In twitter ionic form, amino acids show amphoteric behavior as they react both with acids and bases. Except glycine, all other naturally occurring α-amino acids are optically active, since the α-carbon atom is asymmetric. These exist both in 'D' and 'L' forms. Most naturally occurring amino acids have L configuration. L-Amino acids are represented by writing the –NH2 group on left hand side.

Non-proteinogenic amino acids:

Amino acids are multifunctional organic compounds that contain at least one amino and one carboxyl group attached to a central carbon atom, whose side chains may vary in length and branching as well as in content of other functional groups or aromatic rings. Amino acids may form numerous molecular structures, where the relative position of the amino and carboxyl function allows their general classification as 2-, 3-, 4- etc. (also referred to as α, β, γ, etc.) amino acids. Most amino acids have at least one asymmetric carbon and are chiral. Amino acids are classified as non-protein when they are not part of the 22 such molecules that are translated into proteins by the standard genetic code. Aside from the twenty standard amino acids and the two special amino acids, there are a vast number of "Non-proteinogenic amino acids (NPA)". Two of these can be encoded in the genetic code, but are rather rare in proteins. Selenocysteine is incorporated into some proteins at a UGA codon, which is normally a stop codon. Pyrrolysine is used by some methanogenic bacteria in enzymes that they use to produce methane. It is coded for with the codon UAG (Krzycki, 2005).



The amino acid selenocysteine

Examples of nonstandard amino acids that are not found in proteins include lanthionine, 2-aminoisobutyric acid, dehydroalanine and the neurotransmitter gamma-aminobutyric acid. Nonstandard amino acids often occur as intermediates in the metabolic pathways for standard amino acids - for example ornithine and citrulline occur in the urea cycle, part of amino acid catabolism (Curis et al., 2005). Nonstandard amino

acids are usually formed through modifications to standard amino acids. For example, homocysteine is formed by the transsulfuration pathway from cysteine or as an intermediate in S-adenosyl methionine metabolism (Brosnan and Brosnan, 2006).

Of the thousands of known non-protein amino acids (NPA), about 300 occur in plants. They are found mostly in a small number of families, such as the Leguminosae, Cucurbitaceae, Sapindacae, Aceraceae and Hippocastenaceae. Many of these NPA are structurally similar to the components of common proteins. The incorporation of NPAs into proteins may be associated with autoimmune diseases in humans. Furthermore, there is evidence of a phenotypic conversion of ras-transformed human cells to normal due to incorporation of the tyrosine analogue, azatyrosine, into cellular protein. One NPA that has received some attention is canavanine, (L-2-amino-4-(guanidinooxy) butyric acid), the guanidinooxy structural analogue of arginine.

Role of Non-proteinogenic amino acids

In cells, especially autotrophs, several non-proteinogenic amino acids are found as metabolic intermediates. However, despite the catalytic flexibility of PLP-binding enzymes, many amino acids are synthesised as keto-acids (e.g. 4-methyl-2-oxopentanoate to leucine) and aminated in the last step, thus keeping the number of non-proteinogenic amino acid intermediates fairly low. Ornithine and citrulline occur in the urea cycle, part of amino acid catabolism (Curis et al., 2005). In addition to primary metabolism, several non-proteinogenic amino acids are precursors or the final production in secondary metabolism to make compounds such as toxins.

Non-protein amino acids have been implicated in plant defense against insect pests. Direct toxic effects occur through interference with animal amino acid metabolism. Nitrogen is stored in a form that is

metabolically inaccessible to herbivores. Some specialist herbivores have specific detoxification mechanisms.

Classification of Proteins:

Proteins are classified on the basis of their chemical composition, shape and solubility into two major categories as discussed below.

(i) Simple proteins:

Simple proteins are those which, on hydrolysis, give only amino acids. According to their solubility, the simple proteins are further divided into two major groups' fibrous and globular proteins.

(a) Fibrous Proteins: These are water insoluble animal proteins eg. collagen (major protein of connective tissues), elastins (protein of arteries and elastic tissues), keratins (proteins of hair, wool, and nails) are good examples of fibrous proteins. Molecules of fibrous proteins are generally long and thread like.

(b) Globular Proteins: These proteins are generally soluble in water, acids, bases or alcohol. Some examples of globular proteins are albumin of eggs, globulin (present in serum), and haemoglobin. Molecules of globular proteins are folded into compact units which are spherical in shape.

(ii) Conjugated proteins:

Conjugated proteins are complex proteins which on hydrolysis yield not only amino acids but also other organic or inorganic components. The non-amino acid portion of a conjugated protein is called prosthetic group.

Unlike simple proteins, conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups. These are

a) Nucleoproteins (protein + nucleic acid)
b) Mucoproteins and glycoprotein's (protein+ carbohydrates)

c) Chromo proteins (proteins + a colored pigment)
d) Lipoproteins (proteins + lipid)
e) Metalloproteinase (metal binding proteins combined with iron, copper or zinc)
f) Phosphoproteins (proteins attached with a phosphoric acid group).

Proteins can also be classified on the basis of functions they perform, as summarized in table 2.

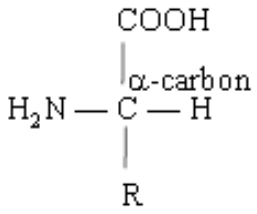
Table 2: Classification of proteins according to their biological functions

Class	Functions	Examples
1. Transport Proteins	Transport of oxygen, glucose and other nutrients	Haemoglobin, Lipoproteins
2.Nutrient and storage Proteins	Store proteins required for the growth of embryo	Gliadin (wheat) Ovalbumin (egg) Casein (milk)
3. Structural Proteins	Give biological structures, strength or protection	Keratin (Hair, nails, etc.) Collagen (cartilage)
4. Defence Proteins	Defend organisms against invasion by other species	Antibodies Snake venoms
5. Enzymes	Act as catalysts in biochemical reactions	Trypsin, Pepsin
6. Regulatory Proteins	Regulate cellular or physiological activity	Insulin

Structure of Proteins

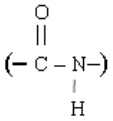
Protein molecules are polymers of different sizes and shapes with different physical and chemical properties. The monomer units for

proteins are amino acids. Al the amino acids that are found in proteins have an amino group (-NH2) on the carbon atom adjacent to carbonyl group, hence are called α-amino acids. The general formula of α-amino acids is shown below.

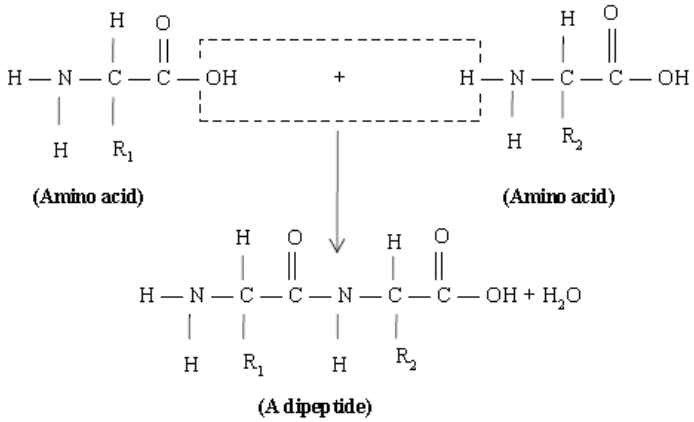


All proteins found in nature are the polymers of about twenty (20) different α-amino acids and these entire have L-configuration. Out of these ten (10) amino acids cannot be synthesized by our body and hence must form the part of our diet. These are called essential amino acids.

All proteins have one common structural feature that heir amino acids are connected to one another by peptide linkages. By a peptide linkage we mean an amide



Bond formed when the carboxyl group of one amino acid molecule reacts with the- amino group of another. In the process, a molecule of water is given of. The product of the reaction is called a peptide or more precisely a dipeptide because it is made by combining two amino acids, as shown below:



If a third amino acid is joined to a dipeptide in the same manner, the product is a tripe tide. Thus, a tripe tide contains three amino acids linked by two peptide linkages. Similar combinations of four, five, six amino acids give a tetra peptide, a pent peptide, a hexapeptide, respectively. Peptides formed by the combination of more than ten amino acid units are called polypeptides. Proteins are polypeptides formed by the combination of large number of amino acid units. There is no clear line of demarcation between polypeptides and proteins. For example insulin, although it contains only 51 amino acids, is generally considered a small protein.

The amino acid unit with the free amino group is known as the N-terminal residue and the one with the free carboxyl group is called the C-terminal residue. By convention, the structure of peptide or proteins written with the N-terminal residue on the left and the C- terminal on the right.

The actual structure of a protein can be discussed at four different levels.

(i) **Primary structure:** Information regarding the sequence of amino acids in a protein chain is called its primary structure. The primary structure of a protein determines its functions and is critical to its biological activity.

(ii) **Secondary structure:** The secondary structure arises due to the regular folding of the polypeptide chain due to hydrogen bonding between and two types of secondary structures have been reported. These are – α helix (Fig.1.3) when the chain coils up and ß-pleated sheet (Fig. 1.4) when hydrogen bonds are formed between the chains.

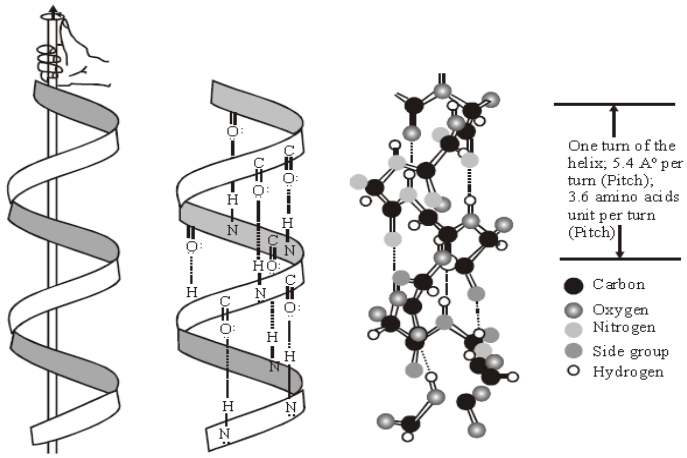
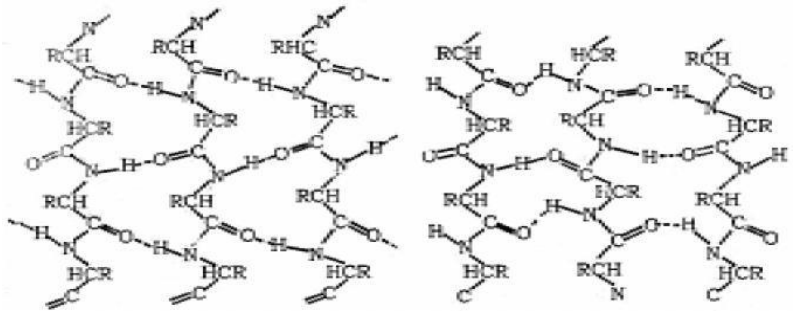


Fig. 1.3: The a-helix structure of protein



Parallel α-Conformation **Ant parallel ß-Conformation**

Fig. 1.4: The ß-pleated-sheet structure of protein

(iii) Tertiary structure: It is the three-dimensional structure of proteins. It arises due to folding and super imposition of various α-helical chains or ß-plated sheets. For example Fig. 1.5 represents the tertiary structure for the protein myoglobin.

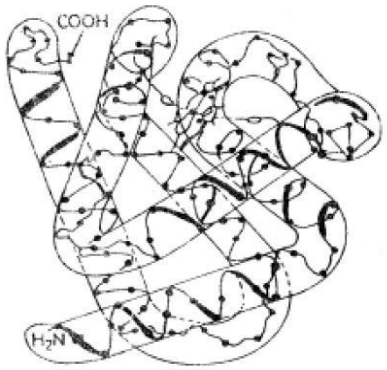


Fig. 1.5: Structure of myoglobin

(iii) **Quaternary structure:** The quaternary structure refers to the way in which simple protein chains associate with each other resulting in

the formation of a complex protein. By different modes of bonding in secondary and tertiary structural levels a protein molecule appears to have a unique three-dimensional structure.

Loss of Protein Structure:

Considering the small differences in the free energy of folded and unfolded proteins, it is not surprising that protein structure is especially sensitive to environmental factors. Many physical and chemical agents can disrupt a protein's native conformation. The process of structure disruption, which may or may not involve protein unfolding, is called denaturation.

Denaturation

One of the great difficulties in the study of the structure of proteins is that if the normal environment of a living protein molecule is changed even slightly, such as by a change in pH or in temperature, the hydrogen bonds are disturbed and broken. When attractions between and within protein molecules are destroyed, the chains separate from each other, globules unfold and helices uncoil. We say that the protein has been denatured. Denaturation is seen in our daily life in many forms. The curdling of milk is caused by bacteria in the milk which produce lactic acid. The change in pH caused by the lactic acid causes denaturation, coagulation and precipitation of the milk proteins. Similarly, the boiling of an egg causes precipitation of the albumin proteins in the egg white. Some proteins (such as those in skin, fingernails, and the stomach lining) are extremely resistant to denaturation.

Biological Importance of Proteins

Of all the molecules encountered in living organisms, proteins have the most diverse functions, as the following list suggests.

1. **Catalysis.** Catalytic proteins called the enzymes accelerate thousands of biochemical reactions in such processes as digestion, energy capture, and biosynthesis. These molecules have remarkable properties. For example, enzymes can increase reaction rates by factors of between 106 and 1012. They can perform this feat under mild conditions of pH and temperature because they can induce or stabilize strained reaction intermediates. For example, ribulose bisphosphate carboxylase is an important enzyme in photosynthesis, and the protein complex nitrogenase is responsible for nitrogen fixation.
2. **Structure.** Structural proteins often have very specialized properties. For example, collagen (the major components of connective tissues) and fibroin (silkworm protein) have significant mechanical strength. Elastin, the rubberlike protein found in elastic fibers, is found in blood vessels and skin that must be elastic to function properly.
3. **Movement.** Proteins are involved in all cell movements. Actin, tubulin, and other proteins comprise the cytoskeleton. Cytoskeletal proteins are active in cell division, endocytosis, exocytosis, and the ameboid movement of white blood cells.
4. **Defense.** A wide variety of proteins are protective. In vertebrates, keratin, a protein found in skin cells, aids in protecting the organism against mechanical and chemical injury. The blood-clotting proteins fibrinogen and thrombin prevent blood loss when blood vessels are damaged. The immunoglobulins (or antibodies) are produced by lymphocytes when foreign organisms such as bacteria invade an organism. Binding antibodies to an invading organism is the first step in its destruction.
5. **Regulation.** Binding a hormone molecule or a growth factor to cognate receptors on its target cell changes cellular function. For example, insulin and glucagon are peptide hormones that regulate

blood glucose levels. Growth hormone stimulates cell growth and division. Growth factors are polypeptides that control animal cell division and differentiation. Examples include platelet-derived growth factor (PDGF) and epidermal growth factor (EGF).

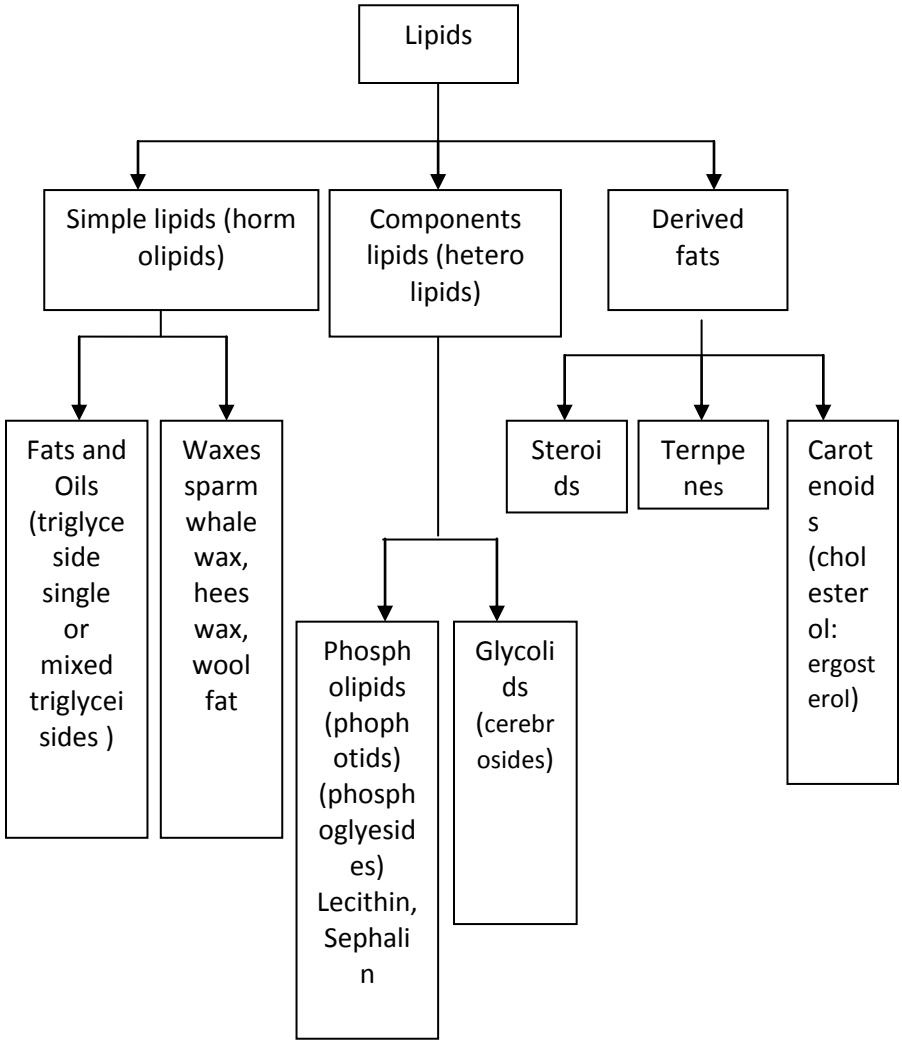
6. **Transport**. Many proteins function as carriers of molecules or ions across membranes or between cells. Examples of membrane transport proteins include the enzyme Na_-K_ ATPase and the glucose transporter. Other transport proteins include hemoglobin, which carries O2 to the tissues from the lungs, and the lipoproteins LDL and HDL, which transport waterinsoluble lipids in the blood from the liver. Transferrin and ceruloplasmin are serum proteins that transport iron and copper, respectively.
7. **Storage.** Certain proteins serve as a reservoir of essential nutrients. For example, ovalbumin in bird eggs and casein in mammalian milk are rich sources of organic nitrogen during development. Plant proteins such as zein perform a similar role in germinating seeds.
8. **Stress response.** The capacity of living organisms to survive a variety of abiotic stresses is mediated by certain proteins. Examples include cytochrome P450, a diverse group of enzymes found in animals and plants that usually convert a variety of toxic organic contaminants into less toxic derivatives, and metallothionein, a cysteine-rich intracellular protein found in virtually all mammalian cells that binds to and sequesters toxic metals such as cadmium, mercury, and silver. Excessively high temperatures and other stresses result in the synthesis of a class of proteins called the heatshock proteins (hsps) that promote the correct refolding of damaged proteins. If such proteins are severely damaged, hsps promote their degradation. (Certain hsps function in the normal process of protein folding). Cells are protected from radiation by DNA repair enzymes.

Lipids

The lipids include a large number of biomolecules of different types. The term lipid originated from a Greek word 'Lipos' meaning fat. In general, those constituents of the cell which are insoluble in water and soluble in organic solvents of low polarity (such as chloroform, ether, benzene etc.) are termed as lipids. Lipids perform a variety of biological functions.

Classification of Lipids

Lipids are classified into three broad categories on the basis of their molecular structure and the hydrolysis products


Lipids
Simple lipids (hormolipids)
Components lipids (heterolipids)
Derived fats
Fats and Oils (triglyceside single or mixed triglyceisides)
Waxes sparm whale wax, hees wax, wool fat
Phospholipids (phophotids) (phosphoglyesides) Lecithin, Sephalin
Glycolids (cerebrosides)
Steroids
Ternpenes
Carotenoids (cholesterol: ergosterol)

(i) **Simple Lipids:** Those lipids which are esters and yield fatty acids and alcohols upon hydrolysis are called simple lipids. They include oils, fats and waxes.

(ii) **Compound Lipids:** Compound lipids are esters of fatty acids and alcohol with additional compounds like phosphoric acid, sugars, proteins etc.

(iii) **Derived Lipids:** Compounds which are formed from oils, fats etc. during metabolism. They include steroids and some fat soluble vitamins.

Structure of lipids

The structure of all three types of lipids is briefly discussed below.

Fatty Acids

Fatty acids consist of a long carbon chain (also called an acyl chain) with carboxylic acid at one end. The vast majority of fatty acids are unbranched linear molecules. The carboxylic acid is ionized at physiological pH (the carboxyl group is deprotonated and therefore negatively charged). Fatty acids in most biological systems are synthesized by serial addition of two carbon units. As a result, fatty acids usually contain an even number of carbons, especially in animals, which synthesize even-chain fatty acids almost exclusively. Biological fatty acids usually contain 14 to 20 carbons, although small amounts of 22 and 24 carbon compounds are found in some tissues. The fatty acid acyl chain "prefers" to be extended, because this results in the least steric hindrance; however, the chain is very flexible, and will adopt a large variety of conformations. The reason for this flexibility is that each carbon-carbon bond can (more or less) freely rotate, and all fatty acids have many carbon-carbon bonds.

Saturated fat and unsaturated fat

Many fatty acids contain double bonds. Fatty acids that do not contain carbon-carbon double bonds are considered to be saturated. When producing unsaturated fatty acids, the biosynthetic machinery incorporates nearly exclusively cis configuration double bonds. Some food products, such as margarine, may contain trans double bonds because of manipulations during food processing; some evidence suggests that this is a potential health problem, and that food products lacking trans fatty acids are therefore healthier. Monounsaturated fatty acids contain a single double bond, while polyunsaturated fatty acids contain more than one site of unsaturation. The cis bond causes a kink in the fatty acid acyl chain (recall that, unlike single bonds, double bonds do not allow free rotation).

"Partially hydrogenated vegetable oil" is a term often found on food ingredient labels. It means that some of the fatty acids in the oil were "hydrogenated" (reduced, so that the double bonds were converted to single bonds). Why is this done? Saturated fat and monounsaturated fat forms a solid at room temperature.

Polyunsaturated fat is (usually) liquid at room temperature. Mixing saturated/monounsaturated/polyunsaturated fatty acids is a way of regulating consistency of food, and also of regulating the consistency of biological systems. This is because, in membranes or in bulk lipid, cis double bonds alter packing density. This decreases van der Waals contacts: the presence of cist double bonds therefore results in lower melting temperature because they result in less regular and less stable structures. The table below shows the effect of chain length and number of cist double bonds on melting temperature (note that the values for melting temperature vary somewhat depending on the reference consulted). Longer chains result in higher melting temperatures;

increasing numbers of double bonds decreases melting temperature for fatty acids of a given length.

Table 3: Saturated Fatty Acids

Number of carbon	Melting point(ºC)
10	32
12	44
14	24
16	63
18	70
20	77
22	82
24	86

Table 4: Effect of Unsaturation on Fatty Acids

Number of carbons	Number of double bonds	Melting Point (°C)
18	0	70
18	1	13
18	2	-5
18	3	-11

Nomenclature

Fatty acids are named using both historical labels and symbols. Both types of nomenclature uniquely define molecules with specific lengths and number and position of double bonds. The table 5 (below)

gives the name and symbol for some of the common physiological fatty acids.

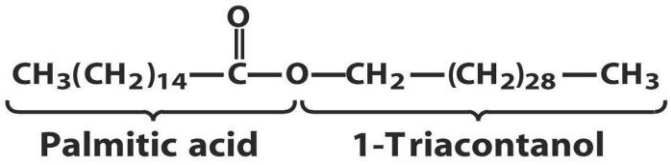
Table 5: Name and symbol for some of the common physiological fatty acids

Number of carbons	Name	Symbol	Structure
12	Lauric acid (dodecanoic acid)	12:0	CH3(CH2)10COOH
14	Myristic acid	14:0	CH3(CH2)12COOH
16	Palmitic acid	16:0	CH3(CH2)14COOH
18	Stearic acid	18:0	CH3(CH2)16COOH
16	Palmitoleic acid	16:1Δ9	CH3(CH2)5CH=CH(CH2)7COOH
18	Oleic acid	18:1Δ9	CH3(CH2)7CH=CH(CH2)7COOH
18	Linoleic acid	18:2Δ9,12	
18	α-Linolenic acid	18:3 Δ9,12,15	
18	γ-Linolenic acid	18:3Δ6,9,12	
20	Arachidonic acid	20:4 Δ5,8,11,14	

Simple Lipids The second types of simple lipids are waxes:

They are the esters of fatty acids with long chain monohydroxy alcohols 26 to 34 carbons atoms. Waxes are wide-spread in nature and occur usually as mixtures. They form a protective coating on the surfaces

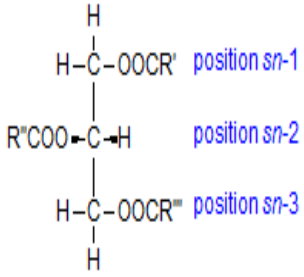
of animals and plants. Some insects also secrete waxes. The main constituent of bees wax obtained from the honey comb of bees is myricyl palmitate:



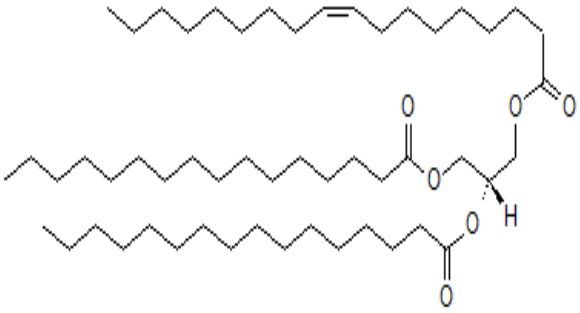
Triacylglycerols:

Triacylglycerol's (formerly called triglycerides) are complex lipids. Triacylglycerols act as energy storage molecules, especially in adipose tissue; triacylglycerols are also found in lipoproteins. Triacylglycerols are not found in membranes, because they are essentially entirely non-polar. Triacylglycerols consist of three fatty acid molecules forming ester links to glycerol. A triacylglycerol molecule can be comprised of different fatty acids, or of three identical fatty acids. In nature, they are synthesized by enzyme systems, which determine that a centre of asymmetry is created about carbon-2 of the glycerol backbone, so they exist in enantiomeric forms, i.e. with different fatty acids in each position.

A stereo specific numbering system has been recommended to describe these forms. In a Fischer projection of a natural L-glycerol derivative, the secondary hydroxyl group is shown to the left of C-2; the carbon atom above this then becomes C-1 and that below is C-3. The prefix "sn" is placed before the stem name of the compound, when the stereochemistry is defined. Their primary biological function is to serve as a store of energy. As an example, the single molecular species 1,2-dihexadecanoyl-3-(9Z-octadecenoyl)-sn-glycerol is illustrated.

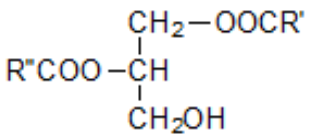


Fischer projection of a triacyl-*sn*-glycerol

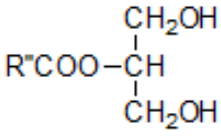


1,2-dihexadecanoyl-3-(9Z-octadecenyl)-*sn*-glycerol

Diacylglycerols (less accurately termed "diglycerides") and monoacylglycerols (monoglycerides) contain two moles and one mole of fatty acids per mole of glycerol, respectively, and exist in various isomeric forms. They are sometimes termed collectively "partial glycerides". Although they are rarely present at greater than trace levels in fresh animal and plant tissues, 1,2-diacyl-sn-glycerols are key intermediates in the biosynthesis of triacylglycerols and other lipids, and they are vital cellular messengers, generated on hydrolysis of phosphatidylinositol and related lipids by a specific phospholipase C. 2-Monoacyl-sn-glycerols are formed as intermediates or end-products of the enzymatic hydrolysis of triacylglycerols; these and other positional isomers are powerful surfactants. 2-Arachidonoylglycerol has important biological properties (as an endocannabinoid).



1,2-/2,3-diacylglycerol

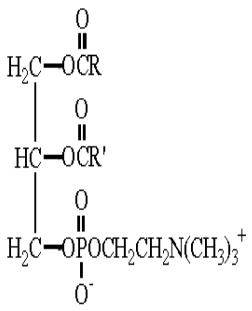


2-monoacylglycerol

Acyl migration occurs rapidly in partial glycerides at room temperature, but especially on heating, in alcoholic solvents or in the presence of acid or base, so special procedures are required for their isolation or analysis if the stereochemistry is to be retained. Synthetic 1/3-monoacylglycerols are important in commerce as surfactants.

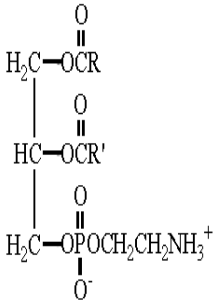
Phospholipids

There are two classes of phospholipids. The first are the glycerophospholipids, which are themselves subdivided into two groups. The first group, phosphatides, is molecules composed of glycerol substituted with two fatty acid esters (just like in fats) and at the third position a phosphate unit connects to an alcohol.



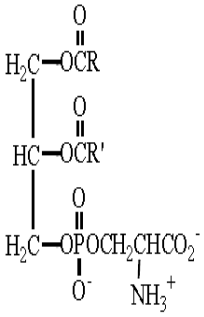
phosphatidylcholine

lecithin



phosphatidylethanolamine

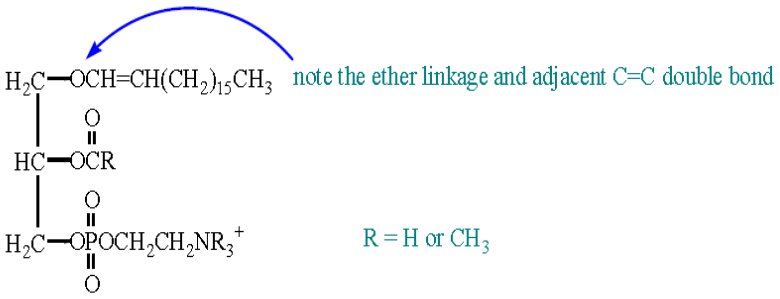
cephalin



phosphatidylserine

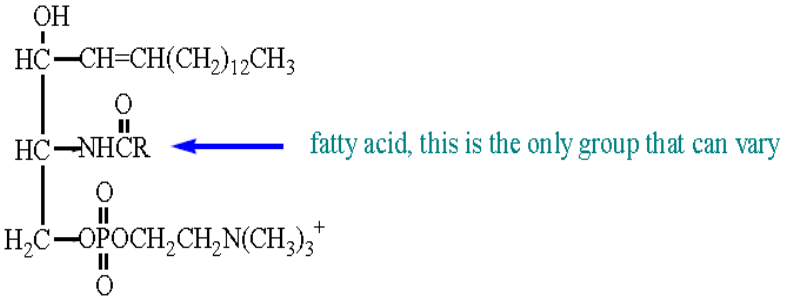
Three alcohols that form phosphatides are choline, ethanolamine, and serine. These compounds are important to the body and are transported as the following phosphatides. The enzymes either cut the molecule free when it is needed or convert it to some other necessary material. The phosphate group and the organic chain attached to it carry electrical charges three phosphatides are components of cell membranes.

Choline is a water-soluble vitamin (as recognized by the Food and Nutrition Board, usually classified as a B vitamin) used to make complex lipids. Phosphatidylcholine is the principal phospholipid of cell membranes. It is also converted to acetyl choline (CH3CO2CH2CH2N (CH3)3+), which is an important neurotransmitter (it carries electrical charges from one nerve cell to another). Choline helps break down homocysteine – a cardiovascular disease risk factor. A lack of this vitamin leads to fatty livers and/or hemorrhagic kidney disease. Serine is the parent of a family of amino acids that also includes glycine and cysteine. Enzymes convert serine (as part of phosphatidylserine) to glycine and cysteine. Serine is also involved in the generation of ethanolamine, which is in turn converted to choline. Interestingly, phosphatidylethanolamine is deficient in Alzheimer's patients. They also act as a histamine blocker in the body. The other subclass of glycerophospholipids is the plasmalogens. These differ from triacylglycerols by even more than the phosphatides. A generic plasmalogen would look like:



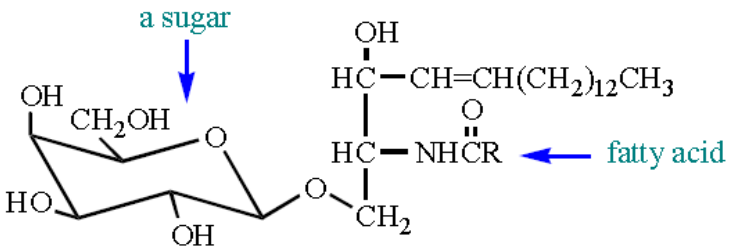
The compound with R = CH3 is called platelet activating factor. It is a strong bronchoconstrictor. It also stimulates other cells to increase their functional and metabolic activities.

The second major class of phospholipids is the sphingolipids. Sphingolipids include the sphingomyelins and cerebrosides. Both are based on the molecule sphingosine. Sphingomyelins have the basic formula:



As the name suggests this lipid is affiliated with the myelin sheath surrounding the cells of the central nervous system. Sphingomyelins comprise about 25% of the lipids in the myelin sheath and their role is key to brain function and electrical transmission through

our nervous system. The other type of sphingolipids we are concerned with are cerebrosides, which are not phospholipids. These compounds are again based on attachments to a sphingosine molecule.

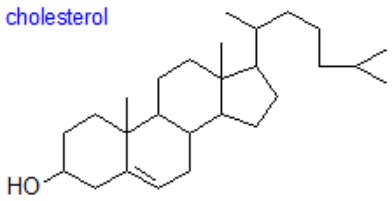


Not surprisingly, these molecules are called glycolipids (cf. glycosides are acetals of sugars). Most of these molecules incorporate ß-D-galactose sugars. Cerebrosides are found most commonly in cell membranes in the brain. One cerebroside found outside the brain, a glucocerebroside, is found in the membranes of macrophages (cells that destroy foreign microorganisms). Several disorders are associated with malfunctioning of sphingolipid metabolism. Probably the best known is Tay-Sachs disease, which strikes infants and is typically fatal by age 3. Niemann-Pick disease also strikes infants and is fatal early in life. Gaucher's disease and Fabry's disease strike later in life and are generally less devastating.

Sterols and sterol esters:

Cholesterol is by far the most common member of a group of steroids in animal tissues; it has a tetra cyclic ring system with a double bond in one of the rings and one free hydroxyl group. It is found both in the Free State, where it has an essential role in maintaining membrane fluidity, and in esterifies form, i.e. as cholesterol esters. Other sterols are present in free and esterifies form in animal tissues, but at trace levels

only. Cholesterol is the precursor of the bile acids and steroidal hormones. In plants, cholesterol is rarely present in other than small amounts, but such phytosterols as sit sterol, stigma sterol, avenasterol, camp sterol and brassicasterol, and their fatty acid esters are usually found, and they perform a similar function. Hopanoids are related lipids produced by some bacterial species.



Membranes

Lipids by definition are water insoluble; however, under certain conditions lipids and water are in fact miscible. Consider a typical water-insoluble fatty acid. At elevated pH values, the fatty acid forms soaps with ions such as Na+ and K+. Fatty acids consist of a polar head and a hydrocarbon or no polar tail. Thus, salts of fatty acids are soluble in both polar and no polar solvents. These types of substances are called amphiphatic compounds, i.e., compounds with both polar and nonpolar components. Amphiphatic substances form a number of different structures. Some of these structures are monolayers, micells, and bilayers.

1. A monolayer of lipid may form at the water–lipid interface (Fig. 1.6)

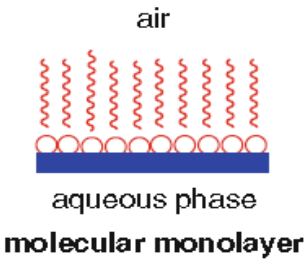


Fig. 1.6: Lipid monolayer illustrating the distribution of the fatty acids between the aqueous and gas (air) phases

2. Soaps and detergents may form structures known as micelles when they reach a defined concentration in solution. This concentration of amphiphatic compound, known as the critical micelle concentration (CMC), is required before micelle formation can occur. The compounds that form micelles are made up of fatty acids with a single hydrophobic tail (Fig. 1.7)

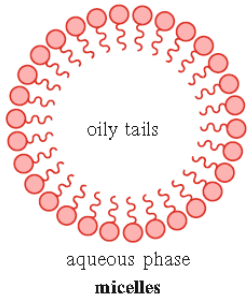


Fig. 1.7: The structure of a typical micelle illustrating the relationship between the fatty acids and the aqueous solution

3. Lipid bilayers may form from lipids that typically contain two hydrophobic tails. The most prominent members of this class are the sphingolipids and glycerophospholipids (Fig. 1.8).

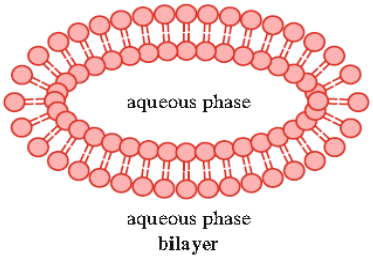


Fig. 1.8: The structure of a synthetic bilayer

Studies have indicated that the bilayer thickness is approximately 60 A °. Lipid bilayers are structurally very similar to biological membranes and thus their properties have been studied extensively. Experimental bilayers, known as liposomes, have been prepared from phospholipids and sphingolipids by sonication.

4. Biological membranes will not be discussed in detail; however, simply stated, they are similar to liposomes in that they contain a lipid bilayer, but unlike liposomes they contain two types of proteins. One, the integral proteins, are embedded in the bilayer and the other, the peripheral proteins, are associated either with the surface of the bilayer or with the integral protein itself. Miscellaneous lipids such as cholesterol are also components of biological membranes and affect its fluidity (Fig. 1.9).

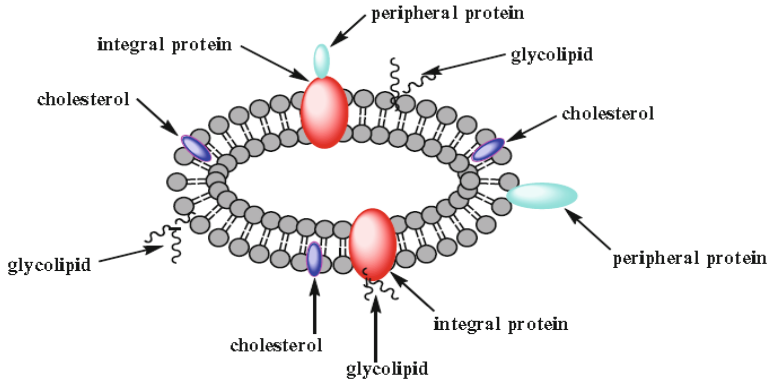


Fig. 1.9: Cartoon of a cell membrane and its many components. The basic structure of the cell membrane is the lipid bilayer

References

1. Krzycki, J. (2005). The direct genetic encoding of pyrrolysine. *Curr Opin Microbiol*., 8 (6): 706-712.
2. Curis E, Nicolis I, Moinard C, Osowska S, Zerrouk N, Bénazeth S, Cynober L (2005). Almost all about citrulline in mammals. *Amino Acids*, 29 (3): 177-205.
3. Brosnan, J. and Brosnan, M. (2006). The sulfur-containing amino acids: an overview. *J Nutr*., 136 (6 Suppl):1636S-1640S.

Suggested Reading

1. Berg, J., J. Tymoczko and L. Stryer. (2010). *Biochemistry: A Short Course*. W. H. Freeman, New York.
2. Donald, V. and Judith, G.V. (2004). *Principles of Biochemistry.* 4th edition. J Wiley & Sons. New York.
3. Lodish, H., Berk, A. Kaiser, C., Krieger, M., Scott, M., Bretscher, A., Plough, H. and Matsudaira, P. (2008). *Molecular Cell Biology*, 6th Edition. W. H. Freeman, New York.
4. Nelson, D. and M. Cox. (2009). *Lehninger Principles of Biochemistry*, 5th Edition. W. H. Freeman, New York.
5. Robert K. M., Darryl K. G., Peter A.M. (2003). *Harper's Illustrated Biochemistry*. 26th edition. McGraw-Hill Medical Publishers. US.
6. Stryer, L. (1995). *Biochemistry;* 4th edition. W. H. Freeman & Company. New York.
7. Watson, J., R. Myers, A. Caudy, and J. Witkowski. (2007). *Recombinant DNA: Genes and Genomes*. W.H. Freeman, New York.
8. Whitford, D. 2005. *Proteins: Structure and Function.* Wiley, New York.

Chapter 2
Computer Applications and Biostatistics

Sandeep Telkar, Nalina. M, Sowmya. H.D, Prakash M Navale,
and Chandrashekara. K.N

Computer application in biology is a complex blend of two distinct scientific disciplines – computer technology and life science. The field of biology is invariably depended on statistics too giving rise to biostatistics. The amalgam of computer applications and biostatistics in combination with several scientific fields gave birth to an interdisciplinary field called bioinformatics, which reaches the biological predictions in an *in silico* way in combination with statistics. Advancement of computer technology has led a path for biological researches giant leap. Progress in the basic functionalities of a computer like store, process, retrieve and reuse, has led to well a synchronized interplay of biology, computing technology and statistics. Biological data is extensive and heterogeneous ranging from text based genome sequences, geometric and spatial information to patterns, large images and simulation. Such a magnitude of information has to be stored, processed, retrieved and reused using proper application software. Internet development, on the other side has catalyzed the interplay by transmission of information. Combination of biostatistics has also added and amplified the elemental aspects of the biological fields like genomics and proteomics, which basically generate enormous amounts of redundant data. Applying computer-intensive biostatistician methods has enabled to process the biological data into information.

Computer Applications and Biostatistics

A computer is an electronic device that can be programmed to carry out a set of arithmetic or logical operations automatically. It is used for almost all the general purpose operations in daily life, since it can perform a sequence of operations, solving more than one kind of problem persistently. Information technologies computational tools and approaches provide opportunities to understand biology in a better way. The information technology addresses several problems of biology like

a) How the biological information can be organized, stored, archived, shared and visualized.
b) How do bimolecular, cells, their populations and other complex biological systems behave under a variety of in vivo conditions?
c) How bimolecular and complex biological process synchronize with each other by mimicking the entire system using modelling and simulation techniques.

And in all the above stated problems, statistics also caters a lot to solve the biological problems and to design algorithms and biometric models, hence the field biostatistics emerges. Computer Applications in biology is known to be computational biology or bioinformatics, which in combination with biostatistics allows one to develop and expand their skills in data management, statistical analyses and representation of resulted data for real world practices. This chapter is focused for introductory concepts of computer, biostatistical theory and their collective application in biology.

1.1 Information and computer technology

The definition of information technology (IT) as defined by the Oxford dictionary is the study and use of electronic system especially with the combination of computers and telecommunications for storing, retrieving, and share information. On contrary the term as described by

free on-line dictionary of computing (FOLDOC), is commonly used as a synonym for computers and computer technology, but it also encompasses other information distribution technologies such as television and telephones. The term is used usually in the context of a business or other enterprise. The term computer technology is usually reserved for the more theoretical, academic aspects of computing.

Computer technology has a great history of development from an ancient digital computation aid 'abacus' to fourth generation core processors. All the developments have been depicted in the table - 1 (Morley, 2014; Parsons, 2011; Jain, 1989; webopedia.com).

Table 1 – Explains all the generations of computers from 1942 to till date. It explains all the key hardware and software technology along with its characteristics and an example for each generation.

Generation of computers	Key hardware technology	Key software technology	Key characteristics	Examples
First Period (1942-1955)	Vacuum tubes; electromagnetic memory	Machine and assembly languages; stored program concept; mostly scientific applications	Bulky in size; highly unreliable limited commercial use; commercial production difficult and costly; difficult to use.	ENIAC,EDVAC,EDSAC,UNI-VAC 1, IBM 701
Second Period (1955-1964)	Transistors magnetic cores memory, magnetic tapes and disks storage	Batch operating System high level programming languages; scientific and commercial applications	Faster, smaller, more reliable, easier and cheaper to produce commercially, easier to use, and easer to upgrade than previous generation systems; scientific; interactive applications.	Honeywell 400,IBM 7030,CDC,1640,UNIVAC

Third (1964-1975)	Integrated circuit (IC) Technology, larger magnetic cores memory; larger capacity disks and magnetic tapes secondary minicomputers	Timesharing operating system standardization of high-level programming languages; unbundling of software from hardware	Faster, Smaller more reliable, easier and cheaper to produce commercially, easier to use and easier to upgrade than previous generation systems; scientific, commercial and interactive online applications	IBM 360/370,PDP-8,PDP-11,CDC 6600
Fourth (1975-1989)	ICs with very large scale integration technology; microprocessors, semiconductor memory, larger capacity hard disks as in-built secondary storage; magnetic tapes and floppy disks as portable storage	Operating systems for PC,GUI, Multiple windows on a single terminal screen; UNIX operating system; C programming language; PC-based and network based applications.	Small affordable reliable and easy to use PCs; more powerful and reliable mainframe systems; totally general purpose machines; easier to produce commercially	IBM PC and its cones, Apple 2,TRS 80, VAX 9000,CRAY-1, CRAY-2,CRAY-X/MP

	media, personal computers, spread of high speed computer networks			
Fifth (1989-Present)	ICs with ultra large scale integration technology larger capacity main memory; larger capacity hard disks; optical disks as portable read-only storage media notebook computers powerful desktop PC's and workstations; very powerful mainframes; internet	World Wide Web; multimedia applications; internet based applications	Portable computers; more powerful, cheaper, reliable and easier to use desktop machines; very powerful mainframes; very high uptime due to hot-pluggable components; totally general purpose machines; easier to produce commercially.	IBM notebooks, Pentium PCs, SUN Workstations, IBM SP/2 SGI Origin 3000, Param 10000, Core processors.

1.1.1 Computer hardware, software and Operating system

Computer Hardware and software are body and soul of a computer. Though these two are entirely different based on physical existence, they have a great interplay and are interdependent. For example, a hardware constitutes a key board for inputting the characters, but is not worth unless until we have a proper application software to be used for typing and vice versa.

- **Computer hardware:-** Computer hardware is the collection of physical entities that constitutes a computer system. It is sensible for our sense organs i.e. it can be seen and are tangible (microsoft.com). Examples include various input and output devices like mouse, keyboard, monitor, graphic cards, sound cards, memory, motherboard, internet patch card, scanner, printer, etc. It also includes important devices like processor, random access memory and hard drive disk (HDD)
- **Computer software:-** Computer software or application software or simply software is a set of instructions that are machine-readable and directs a computer's processor to perform specific operations. It is not sensible for our sense organs i.e. it can't be seen and is intangible (microsoft.com).
- **Operating system: -**The operating system (OS) of the computer is also a kind of software that manages computer hardware and other software resources. It acts as the facilitator between the user and the computer (Silberschatz et al., 2013).

Golftheman in 2010 has better described in his work, the relation between hardware, software, operating system and a computer user. Figure-1 shows how these are interrelated to each other.

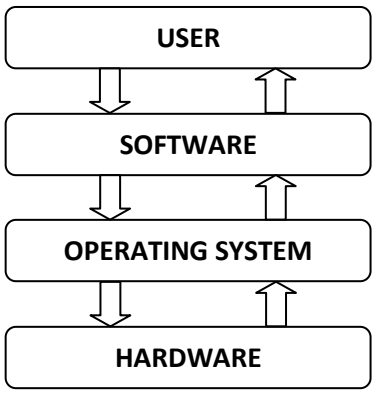


Fig. 1: the figure represents inter-relation between the user, software, OS and hardware.

1.2 Biostatistics

Statistics is a combination of logic and mathematics. Biostatistics is the science, which utilizes mathematical logic to the analysis and interpretation of the biological information and observations. Unlike the other scientific disciplines, Statistics is not a body of substantive knowledge, but only a body of methods of obtaining knowledge. It deals with only numerical data. Information supported by the numerical facts is precise and hence statistical information is more meaningful and measurable than the non-measurable information. For example, the statement “There is 90% chance of occurring of an event” is more accurate than the statement “There is a very good chance of occurring of the event”. In the ordinary sense the term 'Statistics' is used for numerical figures or statistical data, whereas in the wider sense the term refers to various statistical methods. Statistics is as old as human society itself. Its origin can be traced back to pre-Christian era. The Pharos (rulers of ancient Egypt) and Hebrews (ancient Palestinians) used to keep a regular record of manpower and material strengths for their military and fiscal (taxation) purposes. Therefore statistics is called the "Science of the Kings". The word 'Statistics' is derived from the (Latin) word “States'” indicating the practice of data gathering by the State/Rulers. The development of the Statistics is also helped by the

gamblers of the seventeenth century. The laws of the chance are very essential to the gamblers to win the games and hence they invented batter methods to find out the probabilities.

Statistical methods are important to draw valid conclusions from the obtained data. This chapter provides background information related to fundamental methods and techniques in biostatistics to analyze and interpret their study data and to critically interpret published literature. Acquiring such skills currently forms an integral part of researcher/students training. It has been commonly seen that students have an inherent apprehension and prefer staying away from biostatistics. Statistics implies both, data and statistical methods. Statistics without scientific application has no roots. Thus, statistics may be defined as the discipline concerned with the treatment of numerical data derived from group of individuals. These individuals may be human beings, animals, or other organisms. Biostatistics is a branch of statistics applied to biological. Biostatistics covers applications and contributions not only from health, medicines and, nutrition but also from fields such as genetics, biology, epidemiology, and many others. Biostatistics mainly consists of various steps like generation of hypothesis, collection of data, and application of statistical analysis. To begin with, readers should know about the data obtained during the experiment, its distribution, and its analysis to draw a valid conclusion from the experiment. Statistical method has two major branches mainly descriptive and inferential. Descriptive statistics explain the distribution of population measurements by providing types of data, estimates of central tendency (mean, mode and median), and measures of variability (standard deviation, correlation coefficient), whereas inferential statistics is used to express the level of certainty about estimates and includes hypothesis testing, standard error of mean and confidence interval.

Types of Data:-

Observations recorded during research constitute data. There are three types of data i.e. nominal, ordinal, and interval data. Statistical methods for analysis mainly depend on type of data.

1) **Nominal data:** This is synonymous with categorical data where data is simply assigned “names” or categories based on the presence or absence of certain attributes/characteristics without any ranking between the categories. For example,

bacterial culture studies are categorized by growth as positive or negative to particular growth media. It also includes binominal data, which refers to two possible outcomes. For example, outcome of cancer may be death or survival, drug therapy with drug 'X' will show improvement or no improvement at all.

2) **Ordinal data**: It is also called as ordered, categorical, or graded data. Generally, this type of data is expressed as scores or ranks. There is a natural order among categories, and they can be ranked or arranged in order. For example, speed may be classified as slow, medium, and fast. Since there is an order between the three grades of speed, this type of data is called as ordinal. To indicate the intensity of speed, it may also be expressed as scores (slow = 1, medium = 2, fast = 3). Hence, data can be arranged in an order and rank.
3) **Interval data**: This type of data is characterized by an equal and definite interval between two measurements. For example, weight is expressed as 20, 21, 22, 23, 24 kg. The interval between 20 and 21 is same as that between 23 and 24. Interval type of data can be either continuous or discrete. A continuous variable can take any value within a given range. For example: hemoglobin (Hb) level may be taken as 11.3, 12.6, 13.4 gm % while a discrete variable is usually assigned integer values i.e. does not have fractional values. For example, number of meals per day by a person is generally discrete variables. Sometimes, certain data may be converted from one form to another form to reduce skewness and make it to follow the normal distribution. For example, plant growth are converted to their log values and plotted in growth response curve to obtain a straight line so that analysis becomes easy. Data can be transformed by taking the logarithm, square root, or reciprocal. Logarithmic conversion is the most common data transformation used in agricultural research.

Measures of Central Tendencies

Mean, median, and mode are the three measures of central tendencies. Mean is the common measure of central

tendency, most widely used in calculations of averages. It is least affected by sampling fluctuations. The mean of a number of individual values (X) is always nearer the true value of the individual value itself. Mean shows less variation than that of individual values, hence they give confidence in using them. It is calculated by adding up the individual values (Σx) and dividing the sum by number of items (n). Suppose height of 7 children's is 60, 70, 80, 90, 90, 100, and 110 cms. Addition of height of 7 children is 600 cm, so mean (X) = Σx/n = 600/7 = 85.71. Median is an average, which is obtained by getting middle values of a set of data arranged or ordered from lowest to the highest (or vice versa). In this process, 50% of the population has the value smaller than and 50% of samples have the value larger than median. It is used for scores and ranks. Median is a better indicator of central value when one or more of the lowest or the highest observations are wide apart or are not evenly distributed. Median in case of even number of observations is taken arbitrary as an average of two middle values, and in case of odd number, the central value forms the median. In above example, median would be 90. Mode is the most frequent value, or it is the point of maximum concentration. Most fashionable number, which occurred repeatedly, contributes mode in a distribution of quantitative data. In above example, mode is 90. Mode is used when the values are widely varying and is rarely used in medical studies. For skewed distribution or samples where there is wide variation, mode, and median are useful. Even after calculating the mean, it is necessary to have some index of variability among the data. Range or the lowest and the highest values can be given, but this is not very useful if one of these extreme values is far off from the rest. At the same time, it does not tell how the observations are scattered around the mean. Therefore, following indices of variability play a key role in biostatistics.

Standard Deviation:

In addition to the mean, the degree of variability of responses has to be indicated since the same mean may be obtained from different sets of values. Standard deviation (SD) describes the variability of the observation about the mean. To

describe the scatter of the population, most useful measure of variability is SD. Summary measures of variability of individuals (mean, median, and mode) are further needed to be tested for reliability of statistics based on samples from population variability of individual. To calculate the SD, we need its square called variance.

Variance is the average square deviation around the mean and is calculated by

Variance = Σ(x-x-) 2/n OR Σ(x-x-) 2/n-1, now SD = √variance.

SD helps us to predict how far the given value is away from the mean, and therefore, we can predict the coverage of values. SD is more appropriate only if data are normally distributed. If individual observations are clustered around sample mean (M) and are scattered evenly around it, the SD helps to calculate a range that will include a given percentage of observation. For example, if N ≥ 30, the range M ± 2(SD) will include 95% of observation and the range M ±3 (SD) will include 99% of observation. If observations are widely dispersed, central values are less representative of data, hence variance is taken. While reporting mean and SD, better way of representation is 'mean (SD)' rather than 'mean ± SD' to minimize confusion with confidence interval.

Correlation Coefficient:

Correlation is relationship between two variables. It is used to measure the degree of linear relationship between two continuous variables. It is represented by 'r'. In Chi-square test, we do not get the degree of association, but we can know whether they are dependent or independent of each other. Correlation may be due to some direct relationship between two variables. This also may be due to some inherent factors common to both variables. The correlation is expressed in terms of coefficient. The correlation coefficient values are always between -1 and +1. If the variables are not correlated, then correlation coefficient is zero. The maximum value of 1 is obtained if there is a straight line in scatter plot and considered as perfect positive correlation. The association is positive if the values of x-axis and y-axis tend to be high or low together. On the contrary, the

association is negative i.e. -1 if the high y axis values tends to go with low values of x axis and considered as perfect negative correlation. Larger the correlation coefficient, stronger is the association. A weak correlation may be statistically significant if the numbers of observation are large. Correlation between the two variables does not necessarily suggest the cause and effect relationship. It indicates the strength of association for any data in comparable terms as for example, correlation between height and weight, age and height, weight loss and poverty, parity and birth weight, socioeconomic status and hemoglobin. While performing these tests, it requires x and y variables to be normally distributed. It is generally used to form hypothesis and to suggest areas of future research.

Types of Distribution:

Though this universe is full of uncertainty and variability, a large set of experimental/biological observations always tend towards a normal distribution. This unique behavior of data is the key to entire inferential statistics. There are two types of distribution.

1. **Gaussian /normal distribution:** If data is symmetrically distributed on both sides of mean and form a bell-shaped curve in frequency distribution plot, the distribution of data is called normal or Gaussian. The noted statistician professor Gauss developed this, and therefore, it was named after him. The normal curve describes the ideal distribution of continuous values i.e. heart rate, blood sugar level and Hb % level. Whether our data is normally distributed or not, can be checked by putting our raw data of study directly into computer software and applying distribution test. Statistical treatment of data can generate a number of useful measurements, the most important of which are mean and standard deviation of mean. In an ideal Gaussian distribution, the values lying between the points 1 SD below and 1 SD above the mean value (i.e. ± 1 SD) will include 68.27% of all values. The range, mean ± 2 SD includes approximately 95% of values distributed about this mean, excluding 2.5% above and 2.5% below the range [Fig. 2]. **In ideal distribution of the values:** the mean, mode, and median are equal within

population under study. Even if distribution in original population is far from normal, the distribution of sample averages tend to become normal as size of sample increases. This is the single most important reason for the curve of normal distribution. Various methods of analysis are available to make assumptions about normality, including‘t’ test and analysis of variance (ANOVA). In normal distribution, skew is zero. If the difference (mean–median) is positive, the curve is positively skewed and if it is (mean–median) negative, the curve is negatively skewed, and therefore, measure of central tendency differs [Fig. 2].

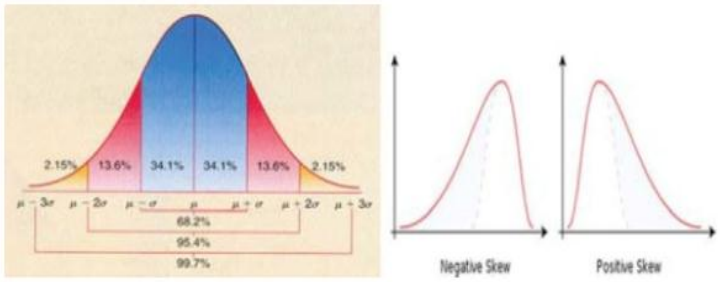


Fig. 2: Diagram showing normal distribution curve with negative and positive skew μ = Mean, σ = Standard deviation

1) **Non-Gaussian (non-normal) distribution:** If the data is skewed on one side, then the distribution is non-normal. It may be binominal distribution or Poisson distribution. In binominal distribution, event can have only one of two possible outcomes such as yes/no, positive/negative, survival/death, and smokers/non-smokers. When distribution of data is non-Gaussian, different test like Wilcoxon, Mann-Whitney, Kruskal-Wallis, and Friedman test can be applied depending on nature of data.

Standard Error of Mean

Since we study some points or events (sample) to draw conclusions about all patients or population and use the sample

mean (M) as an estimate of the population mean (M1), we need to know how far M can vary from M1 if repeated samples of size N are taken. A measure of this variability is provided by Standard error of mean (SEM), which is calculated as (SEM = SD/√n). SEM is always less than SD. What SD is to the sample, the SEM is to the population mean.

Applications of Standard Error of Mean:

Applications of SEM include:

1) To determine whether a sample is drawn from same population or not when it's mean is known.
2) To work out the limits of desired confidence within which the population mean should lie. For example, take fasting blood sugar of 200 lawyers. Suppose mean is 90 mg% and SD = 8 mg%. With 95% confidence limits, fasting blood sugar of lawyer's would be; n = 200, SD = 8; hence SEM = SD/√n=8/√200=8/14.14=0.56. Hence, Mean fasting blood sugar + 2 SEM = 90 + (2 × 0.56) = 91.12 while Mean fasting blood sugar - 2 SEM = 90 - (2 × 0.56) = 88.88. So, confidence limits of fasting blood sugar of lawyer's population are 88.88 to 91.12 mg %. If mean fasting blood sugar of another lawyer is 80, we can say that, he is not from the same population.

Confidence Interval (CI):

Confidence limits are two extremes of a measurement within which 95% observations would lie. These describe the limits within which 95% of the mean values if determined in similar experiments are likely to fall. The value of 't' corresponding to a probability of 0.05 for the appropriate degree of freedom is read from the table of distribution. By multiplying this value with the standard error, the 95% confidence limits for the mean are obtained as per formula below.

Lower confidence limit = mean - (t0.05 × SEM) Upper confidence limit = mean + (t0.05 × SEM) If n > 30, the interval M ± 2(SEM) will include M with a probability of 95% and the interval M ± 2.8 (SEM) will include M with probability of 99%. These intervals are, therefore, called the 95% and 99% confidence intervals, respectively. The important difference between the 'p'

value and confidence interval is that confidence interval represents clinical significance, whereas 'p' value indicates statistical significance. Therefore, in many clinical studies, confidence interval is preferred instead of 'p' value, and some journals specifically ask for these values. Various medical journals use mean and SEM to describe variability within the sample. The SEM is a measure of precision for estimated population mean, whereas SD is a measure of data variability around mean of a sample of population. Hence, SEM is not a descriptive statistics and should not be used as such. Correct use of SEM would be only to indicate precision of estimated mean of population.

Null Hypothesis:

The primary object of statistical analysis is to find out whether the effect produced by a compound under study is genuine and is not due to chance. Hence, the analysis usually attaches a test of statistical significance. First step in such a test is to state the null hypothesis. In null hypothesis (statistical hypothesis), we make assumption that there exist no differences between the two groups. Alternative hypothesis (research hypothesis) states that there is a difference between two groups. For example, a new drug 'A' is claimed to have analgesic activity and we want to test it with the placebo. In this study, the null hypothesis would be 'drug A is not better than the placebo.' Alternative hypothesis would be 'there is a difference between new drug 'A' and placebo.' When the null hypothesis is accepted, the difference between the two groups is not significant. It means, both samples were drawn from single population, and the difference obtained between two groups was due to chance. If alternative hypothesis is proved i.e. null hypothesis is rejected, then the difference between two groups is statistically significant. A difference between drug 'A' and placebo group, which would have arisen by chance is less than five percent of the cases, that is less than 1 in 20 times is considered as statistically significant ($P < 0.05$). In any experimental procedure, there is possibility of occurring two errors.

1) Type I Error (False positive)

This is also known as α error. It is the probability of finding a difference; when no such difference actually exists, which results in the acceptance of an inactive compound as an active compound. Such an error, which is not unusual, may be tolerated because in subsequent trials, the compound will reveal itself as inactive and thus finally rejected. For example, we proved in our trial that new drug 'A' has an analgesic action and accepted as an analgesic. If we commit type I error in this experiment, then subsequent trial on this compound will automatically reject our claim that drug 'A' is having analgesic action and later on drug 'A' will be thrown out of market. Type I error is actually fixed in advance by choice of the level of significance employed in test. It may be noted that type I error can be made small by changing the level of significance and by increasing the size of sample.

2) Type II Error (False negative)

This is also called as β error. It is the probability of inability to detect the difference when it actually exists, thus resulting in the rejection of an active compound as an inactive. This error is more serious than type I error because once we labeled the compound as inactive, there is possibility that nobody will try it again. Thus, an active compound will be lost. This type of error can be minimized by taking larger sample and by employing sufficient dose of the compound under trial. For example, we claim that drug 'A' is not having analgesic activity after suitable trial. Therefore, drug 'A' will not be tried by any other researcher for its analgesic activity and thus drug 'A', in spite of having analgesic activity, will be lost just because of our type II error. Hence, researcher should be very careful while reporting type II error.

Level of Significance:

If the probability (P) of an event or outcome is high, we say it is not rare or not uncommon. But, if the P is low, we say it is rare or uncommon. In biostatistics, a rare event or outcome is called significant, whereas a non-rare event is called non-significant. The 'P' value at which we regard an event or outcomes as enough to be regarded as significant is called the

significance level. In research, most commonly P value less than 0.05 or 5% is considered as significant level. However, on justifiable grounds, we may adopt a different standard like P < 0.01 or 1%. Whenever possible, it is better to give actual P values instead of P < 0.05. Even if we have found the true value or population value from sample, we cannot be confident as we are dealing with a part of population only; howsoever big the sample may be. We would be wrong in 5% cases only if we place the population value within 95% confidence limits. Significant or insignificant indicates whether a value is likely or unlikely to occur by chance. 'P' indicates probability of relative frequency of occurrence of the difference by chance.

Outliers:

Sometimes, when we analyze the data, one value is very extreme from the others. Such value is referred as outliers. This could be due to two reasons. Firstly, the value obtained may be due to chance; in that case, we should keep that value in final analysis as the value is from the same distribution. Secondly, it may be due to mistake. Causes may be listed as typographical or measurement errors. In such cases, these values should be deleted, to avoid invalid results.

One-tailed and two-tailed Test:

When comparing two groups of continuous data, the null hypothesis is that there is no real difference between the groups (A and B). The alternative hypothesis is that there is a real difference between the groups. This difference could be in either direction e.g. A > B or A < B. When there is some sure way to know in advance that the difference could only be in one direction e.g. A > B and when a good ground considers only one possibility, the test is called one-tailed test. Whenever we consider both the possibilities, the test of significance is known as a two-tailed test. For example, when we know that English boys are taller than Indian boys, the result will lie at one end that is one tail distribution, hence one tail test is used. When we are not absolutely sure of the direction of difference, which is usual, it is always better to use two-tailed test. For example, a new drug 'X'

is supposed to have an antihypertensive activity, and we want to compare it with atenolol. In this case, as we don't know exact direction of effect of drug 'X', so one should prefer two-tailed test. When you want to know the action of particular drug is different from that of another, but the direction is not specific, always use two-tailed test. At present, most of the journals use two-sided P values as a standard norm in biomedical research.

Importance of Sample Size Determination

Sample is a fraction of the universe. Studying the universe is the best parameter. But, when it is possible to achieve the same result by taking fraction of the universe, a sample is taken. Applying this, we are saving time, manpower, cost, and at the same time, increasing efficiency. Hence, an adequate sample size is of prime importance in biomedical studies. If sample size is too small, it will not give us valid results, and validity in such a case is questionable, and therefore, whole study will be a waste. Furthermore, large sample requires more cost and manpower. It is a misuse of money to enroll more subjects than required. A good small sample is much better than a bad large sample. Hence, appropriate sample size will be ethical to produce precise results.

Factors Influencing Sample Size Include

1) Prevalence of particular event or characteristics- If the prevalence is high, small sample can be taken and vice versa. If prevalence is not known, then it can be obtained by a pilot study.
2) Probability level considered for accuracy of estimate- If we need more safeguard about conclusions on data, we need a larger sample. Hence, the size of sample would be larger when the safeguard is 99% than when it is only 95%. If only a small difference is expected and if we need to detect even that small difference, then we need a large sample.
3) Availability of money, material, and manpower.
4) Time bound study curtails the sample size as routinely observed with dissertation work in post graduate courses.

Sample Size Determination and Variance Estimate

To calculate sample size, the formula requires the knowledge of standard deviation or variance, but the population variance is unknown. Therefore, standard deviation has to be estimated. Frequently used sources for estimation of standard deviation are:

i. A pilot or preliminary sample may be drawn from the population, and the variance computed from the sample may be used as an estimate of standard deviation. Observations used in pilot sample may be counted as a part of the final sample.
ii. Estimates of standard deviation may be accessible from the previous or similar studies, but sometimes, they may not be correct.

Calculation of Sample Size

Calculation of sample size plays a key role while doing any research. Before calculation of sample size, following five points are to be considered very carefully. First of all, we have to assess the minimum expected difference between the groups. Then, we have to find out standard deviation of variables. Different methods for determination of standard deviation have already been discussed previously. Now, set the level of significance (alpha level, generally set at $P < 0.05$) and Power of study (1-beta = 80%). After deciding all these parameters, we have to select the formula from computer programs to obtain the sample size. Various softwares are available free of cost for calculation of sample size and power of study. Lastly, appropriate allowances are given for non compliance and dropouts, and this will be the final sample size for each group in study. We will work on two examples to understand sample size calculation. a) The mean (SD) diastolic blood pressure of hypertensive patient after Enalapril therapy is found to be 88(8). It is claimed that Telmisartan is better than Enalapril, and a trial is to be conducted to find out the truth. By our convenience, suppose we take minimum expected difference between the two groups is 6 at significance level of 0.05 with 80% power. Results will be analyzed by unpaired‘t’ test. In this case, minimum expected difference is

6, SD is 8 from previous study, alpha level is 0.05, and power of study is 80%. After putting all these values in computer program, sample size comes out to be 29. If we take allowance to non-compliance and dropout to be 4, then final sample size for each group would be 33.

b) The mean hemoglobin (SD) of newborn is observed to be 10.5 (1.4) in pregnant mother of low socioeconomic group. It was decided to carry out a study to decide whether iron and folic acid supplementation would increase hemoglobin level of newborn. There will be two groups, one with supplementation and other without supplementation. Minimum difference expected between the two groups is taken as 1.0 with 0.05 level of significance and power as 90%. In this example, SD is 1.4 with minimum difference 1.0. After keeping these values in computer-based formula, sample size comes out to be 42 and with allowance of 10%, final sample size would be 46 in each group.

Power of Study

It is a probability that study will reveal a difference between the groups if the difference actually exists. A more powerful study is required to pick up the higher chances of existing differences. Power is calculated by subtracting the beta error from 1. Hence, power is (1-Beta). Power of study is very important while calculation of sample size. Power of study can be calculated after completion of study called as posteriori power calculation. This is very important to know whether study had enough power to pick up the difference if it existed. Any study to be scientifically sound should have at least 80% power. If power of study is less than 80% and the difference between groups is not significant, then we can say that difference between groups could not be detected, rather than any difference between the groups. In this case, power of study is too low to pick up the exiting difference. It means probability of missing the difference is high and hence the study could have missed to detect the difference. If we increase the power of study, then sample size also increases. It is always better to decide power of study at initial level of research.

How to Choose an Appropriate Statistical Test

There are number of tests in biostatistics, but choice mainly depends on characteristics and type of analysis of data. Sometimes, we need to find out the difference between means or medians or association between the variables. Number of groups used in a study may vary; therefore, study design also varies. Hence, in such situation, we will have to make the decision which is more precise while selecting the appropriate test. In appropriate test will lead to invalid conclusions. Statistical tests can be divided into parametric and non- parametric tests. If variables follow normal distribution, data can be subjected to parametric test, and for non- Gaussian distribution, we should apply non-parametric test. Statistical test should be decided at the start of the study. Following are the different parametric test used in analysis of various types of data.

1) Student's 't' Test

Mr. W. S. Gosset, a civil service statistician, introduced't' distribution of small samples and published his work under the pseudonym 'Student.' This is one of the most widely used tests in pharmacological investigations, involving the use of small samples. The't' test is always applied for analysis when the number of sample is 30 or less. It is usually applicable for graded data like blood sugar level, body weight, height etc. If sample size is more than 30, 'Z' test is applied. There are two types of 't' test, paired and unpaired.

When to apply paired and unpaired

a) When comparison has to be made between two measurements in the same subjects after two consecutive treatments, paired 't' test is used. For example, when we want to compare effect of drug A (i.e. decrease blood sugar) before start of treatment (baseline) and after 1 month of treatment with drug A.

b) When comparison is made between two measurements in two different groups, unpaired 't' test is used. For example, when we compare the effects of drug A and B (i.e. mean change in blood sugar) after one month from baseline in both groups, unpaired 't' test' is applicable.

2) ANOVA

When we want to compare two sets of unpaired or paired data, the student's't' test is applied. However, when there are 3 or more sets of data to analyze, we need the help of well-designed and multi-talented method called as analysis of variance (ANOVA). This test compares multiple groups at one time. In ANOVA, we draw assumption that each sample is randomly drawn from the normal population, and also they have same variance as that of population. There are two types of ANOVA.

A) **One way ANOVA:** It compares three or more unmatched groups when the data are categorized in one way. For example, we may compare a control group with three different doses of aspirin in rats. Here, there are four unmatched group of rats. Therefore, we should apply one way ANOVA. We should choose repeated measures ANOVA test when the trial uses matched subjects. For example, effect of supplementation of vitamin C in each subject before, during, and after the treatment. Matching should not be based on the variable you are com paring. For example, if you are comparing blood pressures in two groups, it is better to match based on age or other variables, but it should not be to match based on blood pressure. The term repeated measures applies strictly when you give treatments repeatedly to one subjects. ANOVA works well even if the distribution is only approximately Gaussian. Therefore, these tests are used routinely in many field of science. The P value is calculated from the ANOVA table.

B) **Two ways ANOVA:** Also called two factors ANOVA, determines how a response is affected by two factors. For example, you might measure a response to three different drugs in both men and women. This is a complicated test. Therefore, we think that for postgraduates, this test may not be so useful.

Importance of post hoc test

Post tests are the modification of't' test. They account for multiple comparisons, as well as for the fact that the comparison are interrelated. ANOVA only directs whether there is significant difference between the various groups or not. If the results are

significant, ANOVA does not tell us at what point the difference between various groups subsist. But, post test is capable to pinpoint the exact difference between the different groups of comparison. Therefore, post tests are very useful as far as statistics is concerned. There are five types of post- hoc test namely; Dunnett's, Turkey, Newman-Keuls, Bonferroni, and test for linear trend between mean and column number.

How to select a post test?

1) Select Dunnett's post-hoc test if one column represents control group and we wish to compare all other columns to that control column but not to each other.
2) Select the test for linear trend if the columns are arranged in a natural order (i.e. dose or time) and we want to test whether there is a trend so that values increases (or decreases) as you move from left to right across the columns.
3) Select Bonferroni, Turkey's, or Newman's test if we want to compare all pairs of columns.

Following are the non-parametric tests used for analysis of different types of data.

1) Chi-square test

The Chi-square test is a non-parametric test of proportions. This test is not based on any assumption or distribution of any variable. This test, though different, follows a specific distribution known as Chi-square distribution, which is very useful in research. It is most commonly used when data are in frequencies such as number of responses in two or more categories. This test involves the calculations of a quantity called Chi-square (x2) from Greek letter 'Chi'(x) and pronounced as 'Kye.' It was developed by Karl Pearson.

Applications

a) Test of proportion: This test is used to find the significance of difference in two or more than two proportions.
b) Test of association: The test of association between two events in binomial or multinomial samples is the most important application of the test in statistical methods. It

measures the probabilities of association between two discrete attributes. Two events can often be studied for their association such as smoking and cancer, treatment and outcome of disease, level of cholesterol and coronary heart disease. In these cases, there are two possibilities, either they influence or affect each other or they do not. In other words, you can say that they are dependent or independent of each other. Thus, the test measures the probability (P) or relative frequency of association due to chance and also if two events are associated or dependent on each other. Varieties used are generally dichotomous e.g. improved / not improved. If data are not in that format, investigator can transform data into dichotomous data by specifying above and below limit. Multinomial sample is also useful to find out association between two discrete attributes. For example, to test the association between numbers of cigarettes equal to 10, 11-20, 21-30, and more than 30 smoked per day and the incidence of lung cancer. Since, the table presents joint occurrence of two sets of events, the treatment and outcome of disease, it is called contingency table (Con- together, tangle- to touch).

How to prepare 2 × 2 table

When there are only two samples, each divided into two classes, it is called as four cell or 2 × 2 contingency table. In contingency table, we need to enter the actual number of subjects in each category. We cannot enter fractions or percentage or mean. Most contingency tables have two rows (two groups) and two columns (two possible outcomes). The top row usually represents exposure to a risk factor or treatment, and bottom row is mainly for control. The outcome is entered as column on the right side with the positive outcome as the first column and the negative outcome as the second column. A particular subject or patient can be only in one column but not in both. The following table explains it in more detail: Even if sample size is small (< 30), this test is used by using Yates correction, but frequency in each cell should not be less than 5. Though, Chi-square test tells an association between two events or characters, it does not measure the strength of association. This is the

limitation of this test. It only indicates the probability (P) of occurrence of association by chance. Yate's correction is not applicable to tables larger than 2 X 2. When total number of items in 2 X 2 table is less than 40 or number in any cell is less than 5, Fischer's test is more reliable than the Chi-square test.

2) Wilcoxon-Matched-Pairs Signed-Ranks Test

This is a non-parametric test. This test is used when data are not normally distributed in a paired design. It is also called Wilcoxon-Matched Pair test. It analyses only the difference between the paired measurements for each subject. If P value is small, we can reject the idea that the difference is coincidence and conclude that the populations have different medians.

3) Mann-Whitney test

It is a Student's't' test performed on ranks. For large numbers, it is almost as sensitive as Student's't' test. For small numbers with unknown distribution, this test is more sensitive than Student's't' test. This test is generally used when two unpaired groups are to be compared and the scale is ordinal (i.e. ranks and scores), which are not normally distributed.

4) Friedman test

This is a non-parametric test, which compares three or more paired groups. In this, we have to rank the values in each row from low to high. The goal of using a matched test is to control experimental variability between subjects, thus increasing the power of the test.

5) Kruskal-Wallis test

It is a non-parametric test, which compares three or more unpaired groups. Non-parametric tests are less powerful than parametric tests. Generally, P values tend to be higher, making it harder to detect real differences. Therefore, first of all, try to transform the data. Sometimes, simple transformation will convert non-Gaussian data to a Gaussian distribution. Non-parametric test is considered only if outcome variable is in rank or scale with only a few categories. In this case, population is far from Gaussian or one or few values are off scale, too high, or too low to measure.

Common problems faced by researcher in any trial and how to address them

Whenever any researcher thinks of any experimental or clinical trial, number of queries arises before him/her. To explain some common difficulties, we will take one example and try to solve it. Suppose, we want to perform a clinical trial on effect of supplementation of vitamin C on blood glucose level in patients of type II diabetes mellitus on metformin. Two groups of patients will be involved. One group will receive vitamin C and other placebo.

a) How much should be the sample size?

In such trial, first problem is to find out the sample size. As discussed earlier, sample size can be calculated if we have S.D, minimum expected difference, alpha level, and power of study. S.D. can be taken from the previous study. If the previous study report is not reliable, you can do pilot study on few patients and from that you will get S.D. Minimum expected difference can be decided by investigator, so that the difference would be clinically important. In this case, Vitamin C being an antioxidant, we will take difference between the two groups in blood sugar level to be 15. Minimum level of significance may be taken as 0.05 or with reliable ground we can increase it, and lastly, power of study is taken as 80% or you may increase power of study up to 95%, but in both the situations, sample size will be increased accordingly. After putting all the values in computer software program, we will get sample size for each group.

b) Which test should I apply?

After calculating sample size, next question is to apply suitable statistical test. We can apply parametric or non-parametric test. If data are normally distributed, we should use parametric test otherwise apply non-parametric test. In this trial, we are measuring blood sugar level in both groups after 0, 6, 12 weeks, and if data are normally distributed, then we can apply repeated measure ANOVA in both the groups followed by Turkey's post-hoc test if we want to compare all pairs of column with each other and Dunnet's post-hoc for comparing 0 with 6 or 12 weeks observations only. If we want to see whether supplementation of vitamin C has any effect on blood glucose

level as compared to placebo, then we will have to consider change from baseline i.e. from 0 to 12 weeks in both groups and apply unpaired't' with two-tailed test as directions of result is non-specific. If we are comparing effects only after 12 weeks, then paired't' test can be applied for intra-group comparison and unpaired 't' test for inter-group comparison. If we want to find out any difference between basic demographic data regarding gender ratio in each group, we will have to apply Chi-square test.

c) Is there any correlation between the variable?

To see whether there is any correlation between age and blood sugar level or gender and blood sugar level, we will apply Spearman or Pearson correlation coefficient test, depending on Gaussian or non-Gaussian distribution of data. If you answer all these questions before start of the trial, it becomes painless to conduct research efficiently.

Softwares for Biostatistics

Statistical computations are now made very feasible owing to availability of computers and suitable software programs. Nowadays, computers are mostly used for performing various statistical tests as it is very tedious to perform it manually. Commonly used software's are MS Office Excel, Graph Pad Prism, SPSS, NCSS, Instant, Dataplot, Sigmastat, Graph Pad Instat, Sysstat, Genstat, MINITAB, SAS, STATA, and Sigma Graph Pad. Statistical methods are necessary to draw valid conclusion from the data. The postgraduate students should be aware of different types of data, measures of central tendencies, and different tests commonly used in biostatistics, so that they would be able to apply these tests and analyze the data themselves. This chapter provides background information, and an attempt is made to highlight the basic principles of statistical techniques and methods.

Biological data and its diversity

Biology grew intensively in 20th century after the discovery of deoxyribose nucleic acid (DNA) in 1953. DNA is a long polymer made from repeating units called nucleotides (Saenger, 1984; Alberts, 2002). DNA was first identified and isolated by Friedrich Miescher

and the double helix structure of DNA was first discovered by James Watson and Francis Crick, by using X-ray crystallographic data collected by Rosalind Franklin and Maurice Wilkins (Watson and Crick, 1953). These discoveries threw great challenges, of which some couldn't be solved in the lab by an experiment. For example, the human genome which contains about three billion nucleotides, of which only few is actually genes. Parsing, searching, and organizing the three billion letters of human DNA are problems that computers are uniquely suited to handle.

Data heterogeneity of biology provided an immense challenge at the beginning of 21st century, when it grew more data and information intensive. Biology in this century is was more of managing the variety and complexity of biological data types leading to the inevitable use of computing technology and statistics. Then the biological information came in many forms and types. For instance,

a) ***Sequence information****:* sequence data having in the form of alphabets were made available. The DNA and protein sequencing projects were on going and were generating an enormous data. For example Human genome project, which was the biggest international project, was completed in 2003 reporting approximately 30,000 genes in humans (genome.gov). Also the genome sequences of organisms including yeast, chicken, fruit flies, mice and several bacteria's were sequenced. These days the high-throughput - next generation sequencing (HT-NGS) technology is a field of genomics research of which is most talked about (Pareek et al., 2011). This technology can produce over 100 times more data compared to the earlier and most sophisticated capillary genome sequencers based on the Sanger method, hence amplifying the biological data challenge for processing into information.
b) ***Spatial information****:* actual biological entities from biomolecules, cells to organism and ecosystems, represent spatial information. For example the three dimensional structure of a protein, encompasses the spatial organization of various amino acids in the entire axis. The structures here are deduced from either of the experimental techniques such as X-ray crystallography or nuclear magnetic resonance (NMR).

c) ***Pattern information****:* Within the genome and proteome the biologically interesting entities are characterized forming patterns which are an important part of sequence analysis. For example, the genome contains patterns associated with genes and also there exists patterns within proteins. Sequence patterns that are widespread are of significance for the organism, as in nature no information is developed for waste, having its own meaning and use.

d) ***Geometric information****:* A great deal of biological function and interaction depends on relative shape of the two molecules interacting within the biological system. The shape of the biomolecules interacting directly depends on the geometrical shape of the interacting partners. For example, the "docking" behavior of a ligand with a biomolecule at a potential binding site depends on the three-dimensional configuration of both of the molecules, expressing the significance of molecular structure data.

Apart from the above mentioned forms and types of biological information, it has appears in several other forms. For example, such as of scalar and vector types of some phenomena that vary in space and time periodically. Also some information is in the form of images produced from the electron and optical microscopes including fluorescence used for identification of expressions. Some are in the form of high-dimensional data, such as the information generated from that of systems biology, where a response of a biomolecule considered as a data point under the influence of sever other similar biomolecules is analyzed. Such systems finally exhibit cellular behaviors like secretion, proliferation and action potentials.

In all, the above mentioned types of biological information includes computational and statistical emphasis and inspire the ways of understanding and processing the biological data leading to interpret the underlying mechanisms of biology and its functions.

Application of computer and statistics in biology

There is a vast and broad application of both computer technology and statistics in biology. Few of the important fields of biology are as mentioned below, briefly explaining the application of both computer technology and statistics.

Genomics

Genomics is an attempt to analyze or compare the complete genetic compliment of an organism. Modern high-density experimental studies on various genomes produce huge amounts of data. Interpretation of this data into a biologically meaningful knowledge is nowadays a major challenge. This challenge is answerable only by the implementation of computer technology and statistics. It requires a development of robust analytical methods which are applicable and useful. Various statistics and algorithms are also used for the comparison of sequenced genome. Principles of Genomics synthesize the state-of-the-art statistical methodologies applied to genome study.

Comparative Genomics is a collection of robust protocols for molecular biologists beginning to use comparative genomic analysis tools in a variety of areas. It involves Comparison of human genetics with model organisms such as mice, fruit fly, *E. coli*. Computational approaches to genome comparison have recently become a common research topic in computer science. A public collection of case studies and demonstrations is growing, ranging from whole genome comparisons to gene expression analysis. This has increased the introduction of different ideas, including concepts from systems and control, information theory, strings analysis and data mining. It is anticipated that computational approaches will become and remain a standard topic for research and teaching, while multiple courses will begin training students to be fluent in both topics.

Proteomics

The term **"proteome"** refers to the entire complement of proteins, including the modifications made to a particular set of proteins, produced by an organism or a cellular system. This will vary with time and distinct requirements, such as stresses, that a cell or organism undergoes. The term **"proteomics"** is a large-scale comprehensive study of a specific proteome, including information on protein abundances, their variations and modifications, along with their interacting partners and networks, in order to understand cellular processes. **“Clinical proteomics”** is a sub-discipline of

proteomics that involves the application of proteomic technologies on clinical specimens such as blood.

Structural genomics or structural bioinformatics

It refers to the analysis of macromolecular structure particularly proteins; the main goal of structural genomics is the extension of idea of genomics, to obtain accurate three dimensional structural models for known proteins. Structural genomics proceeds through increasing levels of analytic resolution, starting with the assignment of genes and markers to individual chromosomes, then the mapping of these genes and markers within a chromosome, and finally the preparation of a physical map culminating in sequencing. Structural genomics takes advantage of completed genome sequences in several ways in order to determine protein structures. The gene sequence of the target protein can also be compared to a known sequence and structural information can then be inferred from the known protein's structure. Structural genomics can be used to predict novel protein folds based on other structural data. Structural genomics can also take modeling-based approach that relies on homology between the unknown protein and a solved protein structure.

Drug designing

Drug design is the inventive process of finding new medications based on the knowledge of a biological target. The drug is most commonly an organic small molecule that activates or inhibits the function of a biomolecule such as a protein, which in turn results in a therapeutic benefit to the patient. In the most basic sense, drug design involves the design of small molecules that are complementary in shape and charge to the biomolecular target with which they interact and therefore will bind to it. Drug design frequently but not necessarily relies on computer modelling techniques. This type of modelling is often referred to as computer-aided drug design. Finally, drug design that relies on the knowledge of the three-dimensional structure of the biomolecular target is known as structure-based drug design.

Functional genomics

Functional genomics refers to the development and application of global experimental approaches to assess gene function by making use of the information and reagents provided by structural genomics. It is characterized by high-throughput or large-scale experimental methodologies combined with statistical or computational analysis of the results (Hieter and Boguski 1997).

Pharmacogenomics

It is an application of genomic approaches and technologies to the identification of drug targets. **In other words** knowing whether a patient carries any of the genetic variations which can help to prescribe and individualize drug therapy, decrease the chance for adverse drug events, and increase the effectiveness of drugs. Pharmacogenomics combines traditional pharmaceutical sciences such as biochemistry with an understanding of common DNA variations in the human genome.

Conclusions

Application of computers and statistics in biology is already widespread. They are extensively and regularly used both for teaching and research. The extensive growth of biology and data produced by its growth is imperative. Biology has more data today, than yesterday and much more data is expected tomorrow. Making sense of such growing biological data is challenging. It would have almost been impossible to analyze, understand and interpret such a huge biological data with the absence of computers and statistics. The ability to extract meaningful information from huge biological data sets and build sophisticated models is altering everything from a gene sequence to drug designing. Computers in combination with statistics allow constructing meaningful models that allows to abstract useful knowledge from biological systems and helps to study them in detail and make some predictions on how biological systems would behave. By these facts, it is evident that no discipline than biology is witnessing more real benefits from computer.

References

1. Alberts, B. (2002). *Molecular Biology of the Cell; Fourth Edition*. New York and London: Garland Science.
2. An Overview of the Human Genome Project. (n.d.). Retrieved October 08, 2014, from http://www.genome.gov/12011238
3. Comparative Genomics. (n.d.). Retrieved October 08, 2014, from http://www.genome.gov/11006946.
4. Computers in Biology - Science Education - National Institute of General Medical Sciences. (n.d.). Retrieved October 05, 2014, from http://publications.nigms.nih.gov/order/computers-biology.html
5. Golftheman. (n.d.). File:Operating system placement.svg - Wikimedia Commons. Retrieved October 13, 2014, from http://commons.wikimedia.org/wiki/File:Operating_system_placement.svg
6. Information technology from FOLDOC. (n.d.). Retrieved October 20, 2014, from http://foldoc.org/information+technology
7. Information technology: definition of information technology in Oxford dictionary (British & World English). (n.d.). Retrieved October 31, 2014, from http://www.oxforddictionaries.com/definition/english/information-technology
8. Morley, D., & Parker, C. (2014). *Understanding Computers: Today and Tomorrow, Comprehensive* (p. 752). Cengage Learning. Retrieved from http://books.google.com/books?id=TVw8AwAAQBAJ&pgis=1
9. Parsons, J. J., & Oja, D. (2010). *New Perspectives on Computer Concepts 2011: Comprehensive* (p. 856). Cengage Learning. Retrieved from http://books.google.com/books?id=tgzQQzp0W5wC&pgis=1
10. Parts of a computer - Windows Help. (n.d.). Retrieved October 20, 2014, from

http://windows.microsoft.com/en-us/windows/computer-parts#1TC=windows-7

11. Saenger, W. (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.
12. Silberschatz, A., Galvin, P., & Gagne, G. (2013). *Operating system concepts*. Retrieved from http://www.just.edu.jo/~misaleh/Teaching/cs375/syllabus.doc
13. The Five Generations of Computers Explained - Webopedia. (n.d.). Retrieved October 20, 2014, from http://www.webopedia.com/DidYouKnow/Hardware_Software/FiveGenerations.asp
14. V.K.Jain. (1989). *Computer For Beginners*. Pustak Mahal. Retrieved from http://books.google.com/books?id=SqEtaAyO--cC&pgis=1
15. WATSON, J. D., & CRICK, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, *171*(4356), 737–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13054692
16. Wooley, J. C., Lin, H. S., & Biology, N. R. C. (US) C. on F. at the I. of C. and. (2005). Catalyzing Inquiry at the Interface of Computing and Biology. National Academies Press (US). Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK25462/

Chapter 3
Macromolecules and Analytical Techniques

Raghavendra. A, Nalina. M and
Chandrashekara K.N.

Living organisms are made up of limited number of types of atom viz; carbon, nitrogen, oxygen, phosphorus, sulphur and ions such as sodium, magnesium, chlorine, potassium, calcium and trace elements such as iron, manganese, cobalt, copper, zinc etc. These combine to form various forms of molecules which are building blocks of life. Among which polysaccharides, proteins, lipids and nucleic acids are some of commonly known macromolecules. Many repetitive monomeric units of simpler organic molecules associate to form macromolecules. Carbohydrates act as instant source of energy for living organism whereas lipids act as stored energy. Proteins are involved in functional and structural need of an organism and nucleic acids acts as informational molecules concerned with carrying genetic information and expression of that information. The properties, structure and functional studies of these macromolecules can be revealed by the use of various analytical tools and techniques. The analytical tools have become the basis for many of the exclusive research and as well as in pharmaceutical industry. Spectroscopy is a tool used for the assay, molecular weight determination and structural arrangements of the basic elements of bio molecules. The centrifugation and chromatographic techniques can be used for isolation, separation and to study chemical and physical properties of bio molecules, whereas electrophoretical techniques are extensively used for separation, molecular characterization and studies on

evolutionary modifications of the macromolecules in biological systems or in living organisms.

Introduction

Living organisms are made up of limited number of types of atom viz; carbon, nitrogen, oxygen, phosphorus, sulphur and ions such as sodium, magnesium, chlorine, potassium, calcium and trace elements such as iron manganese, cobalt, copper, zinc etc. Among which Carbon is a basic element found in all living organisms as well as in macromolecules. The importance of carbon atom in living system is influenced by its ability of forming strong covalent bonds with other elements. Sometimes they form chains and rings which forms skeletons of organic molecules and hence of life itself. To these carbon skeletons, groups of other atoms are added which are called functional groups (e.g., hydroxyl, carbonyl, carboxyl, methyl, ethyl, phenyl, disulphide, phosphoryl, ether, thioester, amino group etc) which confer specific chemical properties on the molecule. These combine to form various forms of molecules which are building blocks of life. Many repetitive monomeric units of simpler organic molecules associate to form macromolecules. They are the major constituents of cells. Among which polysaccharides, proteins, lipids and nucleic acids are some of commonly known macromolecules. Their constituent monomers are monosaccharides, amino acids and nucleotides respectively. Carbohydrates act as instant source of energy for living organism whereas lipids act as stored energy. Proteins are involved in functional and structural need of an organism and nucleic acids acts as informational molecules concerned with carrying genetic information and expression of that information The properties, structure and functional studies of these macromolecules can be revealed by the use of various analytical tools and techniques. The analytical tools have become the basis for many of the modern biological research and as well as in chemical

and pharmaceutical industry. Different types of Centrifugation, Spectroscopy, and Chromatography and Electrophoretical techniques are being used for isolation, assay, molecular weight determination, structural arrangements, for study of chemical and physical properties, molecular characterization and studies on evolutionary modifications of the macromolecules in biological systems or in living organisms. In the present chapter a brief description of the macro molecules is given. Even though variety of analytical techniques are available and employed for various purpose, few important techniques along with their applications are illustrated in this chapter.

Macromolecules

Macromolecules are composed of monomeric subunits and these subunits are made up of simple structure. Macromolecules are organic solid matter of the cells and of four types viz: carbohydrates, proteins, lipids and nucleic acids. However the inorganic salts and minerals also constitute a small fraction of the dry weight of the cell.

1. Carbohydrates.

Carbohydrates are polyhydroxy aldehydes or ketones or substances that yield one of these compounds on hydrolysis. They contain hydrogen and oxygen in 2:1 ratio. In plants they will be present in the form of starch and in animals as glycogen. Carbohydrates can be divided into sugars and non sugars. Sugars are crystalline, soluble and most have a sweet taste. Further sugars are divided into mono saccharides and oligosaccharides. Monosaccharide's are polyhydroxy aldehydes or ketones which cannot be further hydrolysed to yield simpler sugars e.g., glucose, fructose and galactose. Monosaccharides are further classified into triose (3C), tetroses (4C), pentoses (5C) and hexoses (6C) depending on their number of carbon atom present. Oligosaccharides are polyhydroxy aldehydes or ketones which yield monosaccharides upon hydrolysis and are classified as Disaccharides (e.g., sucrose and

maltose) and trisaccharides (e.g., raffinose and stachyose). The monomeric units in these saccharides are held by glycosidic bonds. The monosaccharides are also called as reducing sugars since they have free aldehydic or ketonic carbon atom and can reduce Feheling's and Benedict's solution. The Di and tri sugars are non reducing sugars. In open chain form of monosaccharide's one of the carbon atoms is double bonded with an oxygen atom to form a carbonyl group. If the carbonyl group is present at the end of the carbon chain the monosaccharide is an aldose. If the carbonyl group is in any other position then the monosaccharide is a ketose. The non sugars are amorphous, tasteless and less insoluble in water. The non sugars are further divided into homopoly saccharides and heteropoly saccharides. The homo polysaccharides (e.g., starch and inulin) are made of same repetitive monosaccharide where as oligo polysaccharides (e.g., heparin and agar-agar) are made up of two or three different types of monosaccharide units. The sugars combine with other elements or functional groups to perform various functions in the biological systems. The presence and function of the above sugars and non sugars in biological system is depicted in the table 1.

Table: 1. Biological functions of the carbohydrates.

Sugars/non sugars	Functions
Starch	Stored food in plants
Glycogen	Stored food in animals
Inulin	Roots and tubers of plants
Chitin	Exoskeleton of insects and shells of crustaceans
Pectin	In fruit and intracellular spaces of tissues
Hyaluronic acid	Found in synovial fluid
Chondroitin sulphate	In cartilage and connective tissue

Dermatin sulphate	In skin
Heparin 6 sulphate	In mast cells of arterial walls acts as anti thrombin activity
N Acetyl Muramic acid N Acetyl Glutamic acid	In plant cell walls
Amino sugars	In DNA and RNA
Ascorbic acid	In iron metabolism and redox reagent

2. Lipids

Lipids are water insoluble organic substances which are formed by condensation reactions between fatty acids and alchohol. They have a general formula R.COOH where R is hydrogen or groups such as $–CH_3$,-C_2H_5 etc. Fatty acids are aliphatic carboxylic acids and have ester or amide linkage. Most naturally occurring fatty acids have even number of carbon atoms between 14 and 22. Fatty acids with more carbon atoms are less soluble. The carbon and hydrogen atoms form a hydrocarbon tail which is hydrophobic in nature or water hating. As the hydrocarbon chain length increases the boiling and melting point of the fatty acids will be increased. If fatty acids contain one or more double bond then they are called as unsaturated. Fatty acid or lipids lacking these double bonds are called as saturated fatty acids. Saturated fatty acids with less than 10 carbon atoms are liquid at room temperature. Unsaturated fatty acids have low melting temperature than the saturated fatty acids. In case of alcohols moiety, most of the lipids are made up of alchohol glycerol. Glycerol has three hydroxyl groups which condense with fatty acids to form triglyceride. Classification of the lipids was done by Bloor in 1925, and can be classified into simple, compound and derived lipids. Simple lipids upon hydrolysis yield one or more type of fatty acid and alcohols for e.g., fats and wax. Compound lipids upon hydrolysis yield fatty acid, glycerol and sugars or phosphoric acid for e.g., spingolipids. Derived lipids are obtained by hydrolysis of simple

or compound lipids and they do not contain the ester linkage for e.g., sterols and prostaglandins. Fatty acids can be classified into three groups *viz*: depending on number of carbon atom, based on length of hydrocarbon chain and based on nature of fatty acids. Depending on number of cabon atoms, fatty acids can be divided into even chain fatty acids (Having even chain of 2, 4, 6 for e.g., acetic acid and butyric acid) and odd chain fatty acids (having odd chain of 3, 5, 7 for e.g., propionoic acid and valvic acid). Based on length of hydrocarbon chain fatty acids can be divided into short chain (having 2 to 6 carbon atoms), medium chain (having 8 to 14 carbon atoms) and long chain (having 16 to 24 carbon atoms) fatty acids. Based on the nature of fatty acids they can be divided into saturated (e.g., butyric and caproic acid) and unsaturated (e.g., PUFA such as linoic acid, linolenic acid and arachidonic acid) fatty acids. Fatty acids and lipids have many important biological functions in the living organisms. They are stored form of energy. They are present in cell membrane. The vitamins such as A, D, E and K are soluble in fat only. Fatty acids are present in mylin sheath of neurons. They are involved in the synthesis of steroid harmones and they also acts as cushioning to many vital organs. Fatty acids are building blocks of phospholipids and glycolipids which are fuel molecules in the form of tri acyl glycerols present in fat or adipose tissue. Phosphoglycerides are found in membrane and differs from triacyl glycerols in that one of the OH groups of glycerol is esterified to phosphoric acid. Lipids with protein are called as apo lipoproteins and with sugars are called as glycolipids. These are involved in many of the cellular reaction, signalling pathways, sites of biological recognition and as transport system at the cell membrane levels. Lipids act as signals in the form of harmones and cofactors and pigments.

3. Proteins.

Proteins are most abundant macromolecules in the biological system and they exist in various forms specific for carrying out particular biological functions. The first person to work out the

complete amino acid sequence of insulin a protein was Fred Sanger in 1953. Proteins are polymers and amino acids are their monomers. In other words, they are made up of amino acids. Naturally occurring proteins are made up of 20 standard amino acids. Standard amino acids are those for which at least one code will exist in DNA. Besides this non protein amino acids (about 150) also exist in biological system, these are produced as metabolic intermediates which are derived from alpha amino acids (e.g., Dopain and GABA which are derived from tyrosin and glutamine respectively). Some of the rare amino acids are also present in proteins which do not have any codes in DNA but they are formed from some of the common amino acids. For e.g., Hydroxyproline is formed from proline and is present in collagen. Similarly hydroxyllysine is made from lysine and is also present in collagen. Amino acids are classified based on polarity *viz:* non polar, neutral, positively charged and negatively charged amino acids. Plants are able to synthesize the required amino acids from simpler substances whereas animals cannot synthesize some of the required amino acids which are termed as essential amino acids and these essential amino acids has to be supplemented through their diet to meet the requirement. Based on this amino acids can be classified into three types viz: essential (e.g. methionin, threonin, lysine, argenin, trypsin, valin and leucin), non essential (e.g., cystin, serin, aspartic acid, alanin, glutanin and tyrosin) and semi essential (e.g., histidin and arginin). The amino acids in the protein are held together by a peptide bond. Further the protein undergoes typical folding to form a particular shape as a result of ionic bond, disulphide bond, hydrogen bond and hydrophobic interactions. These bonds play essential roles in the primary, secondary, tertiary and quaternary structures of the protein. The peptide bond is formed by linking amino acid through amide bond. The alpha carboxyl group of one amino acid reacts with alpha amino group of another amino acid to form bond with elimination of water molecule to from peptide. Simple peptides containing two, three or four amino acid residues

are called di, tri and tetra peptides respectively. The peptide possesses free amino and carboxyl groups to which other amino acids can be joined to form a polypeptide. A protein may contain several polypeptide chains. Proteins are meant to perform various roles in biological systems. Based on this they can be classified into catalytic (ribinuclease), transport (haemoglobin and myoglobin), nutrient (egg, ovalbumin and casein), contractile and motile (actin and myosin), structural (collagen, alpha keratin and elastin), defence (immunoglobin, thrombin, venom and bacterial toxins) and regulatory (enzymes and membrane transport) proteins. Based on shape, proteins can be classified into two type's globular and fibrous type. Globular proteins are water soluble and most of them have tertiary structure of proteins for e.g., storage and defence proteins. The fibrous proteins are water insoluble and are of secondary structure in nature for e.g., elastin, alpha keratin etc. Three types of proteins are there based upon their solubility *viz;* simple, conjugated and derived protein. Among these proteins conjugated proteins are associated with some of the chemical component along with the amino acids. The non amino acid part of conjugated protein is called as prosthetic group. Conjugated proteins are classified based on the chemical nature of their prosthetic group. For e.g., lipoprotein contain lipids as their prosthetic part (e.g., beta lipoprotein of blood), similarly glycoprotein (e.g., immunoglobulin G) contain carbohydrates, phosphoprotein (casein of milk) contain phosphate group, metalloprotein contain metal ions such as iron (e.g., ferretin), zinc (e.g., alcohol dehydrogenase), copper (e.g., plastocyanin) etc, as their prosthetic group. Protein exist in several types of structures *viz*; primary, secondary, tertiary and quaternary structures. In primary structure of proteins the amino acids are sequentially arranged in a polypeptide chain. The secondary structure is stable arrangements of amino acids giving rise to recurring structural patterns or it is local special arrangements of polypeptide back bone atoms, it excludes confirmation of side chains. Tertiary structure is three dimensional

folding of polypeptide chain. Quaternary structure of protein is special arrangements of subunits present in the protein. Heat, radiation, strong acids and alkalis, heavy metals, organic solvents and detergents can cause temporary or permanent damage to the structures of protein as a result, protein will no longer perform its biological function this is called denaturation. Sometimes proteins fold back to from its original structure after its denaturation and are able to perform its function this is called as renaturation.

4. Nucleic acids

Nucleic acids are biopolymers of high molecular weight with mononuceotides as their repeating units. These units are arranged to form extremely long molecules known as polynuceotides. They form the genetic material for all living organisms. There are two units of nucleic acids they are Deoxy Ribose Nucleic Acids (DNA) and Ribose Nucleic Acids (RNA).Upon hydrolysis they yield phosphoric acid, pentose sugar and a nitrogenous base. In both DNA and RNA, sugar is in furanose form and is of beta configuration. The 5 carbon sugar is present and generally called as pentose sugar. In DNA it is deoxy ribose where an oxygen atom is removed from the carbon atom 2. These sugars form esters with H_3PO_4 forming a 3^1 5^1 phosphodiester bond between adjacent sugars. The nitrogenous base is of four types, among which two are derived from Purine and two from pyrimidine. The Purine bases are adenine (A) and guanine (G). They contain 6 member pyramidin ring fused to a 5 member imidazole ring. The two pyrimidine are thymine (T) and cytosine(C) in DNA, whereas uracil (U) will be present in RNA in place of thymine. The phosphoric acid gives acidic nature to the nucleic acids. Al together phosphate group and sugars perform structural role in nucleic acids. The nitrogenous bases are conjugated with pentose sugars by beta glycosidic linkage without any phosphate group forming nucleosides. Nucleosides further undergo condensation with phosphoric acid to form nucleotides or in other words nucleotides are phosphoric esters of nucleosides. These nucleotides are not only the building blocks of

nucleic acid they also function as carriers of chemical energy, as components of enzyme factors (coenzyme A, NAD and NADP),as chemical messengers (e.g., c amp). The base composition of the DNA was proposed by Erwin Chargaff in 1940.The ratio of A: T and G:C is 1. When Adenine and Thymine is more in DNA then it is called as AT type. If A: T value is less than 1 then DNA is designated as GC type. The features of DNA were well explained by Watson and Crick in 1953. DNA consists of two polynucleotide chains and the two chains coil around each other (in anti parallel direction) to form a right handed double helical structure. Each chain has a sugar –phosphate backbone with bases which projects at right angles and hydrogen bond with bases of opposite chain across the double helix. The Purine and pyrymidine bases of both the strands are stacked inside the double helix. The offset pairing of the two strands create a major groove measuring $12A^0$ and minor groove measuring $6A^0$ on the surface of duplex. Along the axis of the molecules the base pairs are 0.34nm apart hence a complete turn of the double helix contain 10 base pairs or measures 3.4nm. The double helical of DNA are complimentary to each other. Hence adenine is bonded with thymine and guanine with cytosine. Base tilt normal to the helix axis is 6^0. The double helical structure is held together by two forces one is hydrogen bond between complimentary bases and second is base stacking interactions. The DNA structures are wrapped around the nucleoproteins which consist of protomines and histones, which are basic in nature. These two bind by electrostatic force. Histones and DNA in water forms nucleosomes. The Watson and Crick DNA structure is also called as B form DNA; it is the most stable structural form of the DNA. Other forms of DNA are A and Z form. The A form is relatively divide of water, helix is wider and has more base pairs (11) than that of B DNA. Base tilt normal to the helix axis is 20^0. Whereas; Z DNA is having the left handed helical rotation. Base tilt normal to the helix axis is 7^0. There are 12 base pairs per helical turn and the structure appears more slender and elongated.

The second form of nucleic acid is RNA. It involves in the expression of the genetic information from the DNA. In gene expression RNA acts as an intermediary by using the information encoded in DNA to specify the amino acid sequence of a functional protein. RNA is alkali labile and native RNA is single stranded. There are three important types of RNA *viz*: rRNA, tRNA and mRNA. rRNA is insoluble and found in ribosomes. It can be sedimented at 100000g for 120 minutes. It makes up 80% of total RNA of cell and represents 40-60% of total weight of RNA. It is involved in the synthesis of protein. mRNA are synthesized on the surface of the DNA by DNA dependent RNA polymerase and it exist only for short period. Average size of m RNA is 900-1500 nucleotide unit. If mRNA codes for a single protein then it is called as monocistronic. If it codes for more than two proteins then it is called as polycistronic. In eukaryotes most of the m RNA are monocistronics. Another form of RNA is t RNA or transfer RNA. It is involved in the transfer of amino acids to the r RNA site for protein synthesis. It reads the information encoded in the m RNA and transfers the amino acid to a growing polypeptide chain during protein synthesis. A typical t RNA has 5^1 phosphorylated terminus and a 3^1 terminus for amino acid attachment. It has a T psi C loop of 5 base pairs which help in binding tRNA to ribosome. Also it has DHU loop of 3-4 base pairs which acts as recognition site for enzymes. A variable arm and an activation loop of 7 base pairs were also present.

Analytical techniques and applications

The analytical techniques have become the integral part of any research and analysis of biomolecules. The brief description of the analytical techniques is here with discussed.

1. Centrifugation

In 1923 Theodor Svedberg and his student H. Rinde had successfully analyzed large-grained sols by use of centrifugation. Centrifugation is a process used to separate particles or concentrate

materials such as cells, sub cellular organelles, viruses, proteins and nucleic acids suspended in a liquid medium. The object moving in circle at a steady angular velocity will experience a force **F** directed outwards; this is the basis of centrifugation. The centrifugal force generated by a centrifuge is determined by the speed of rotation and the distance of the sample from the axis of rotation. The centrifugal force is usually compared to the force of gravity and is reported as the relative centrifugal force (RCF) in **g** units. (A force of 1 g is equal to the force of gravity at the earth's surface) The following formula relates the RCF to the rotor speed and diameter: **RCF = (1.119 x 10^{-5}) * (rpm)^2 * r**. where **r** is the radius of rotation (the radius of the rotor) in cm and **rpm** is the centrifuge speed in revolutions per minute. The number 1.119 x 10^{-5} is a conversion factor that allows us to use "**rpm**" and "**g**" units. The RCF is dependent upon the radius of the rotor, the speed at which it rotates, and the design of the rotor itself (fixed angle or swinging bucket). Rotor speed and design can be held constant, but the radius will vary from the top of a centrifuge tube to the bottom. If a measurement for the radius is taken as the mid-point, or as an average radius, and all forces are mathematically related to gravity, then one obtains relative centrifugal force, labelled as **g.** The particles are separated with respect to their molecular weight or saturation coefficient. The two common types of centrifugation are analytical and preparative; the distinction between the two is based on the purpose of centrifugation. Analytical centrifugation involves measuring the physical properties of the sedimenting particles such as sedimentation coefficient or molecular weight. Optical methods are used in analytical ultracentrifugation. Molecules are observed by optical system during centrifugation, to allow observation of macromolecules in solution as they move in gravitational field. The other form of centrifugation is called preparative and the objective is to isolate specific particles which can be reused. There are four basic types of centrifuges *viz*:

Clinical centrifuges, Microfuge, High speed centrifuges and Ultracentrifuges.

- **Clinical centrifuges** – Low cost tabletop centrifuges that can run at a maximum speed of 4000-5000 rpm. These can pellet cells, but not organelles or bio molecules. Two types of rotors are used in it, fixed angle and swinging bucket. The separation can be carried out in 10, 50 or 100 ml tubes.
- **Microfuge** - These can typically run at a speed of up to 14,000 rpm, sufficient speed to pellet nucleic acids and denatured proteins. Microfuges can be used for small volumes sample of 1-2ml.
- **High speed centrifuges** – These operate at maximum speed of 25,000 rpm or about 90000g, which is sufficient to pellet cell nuclei and most bio molecules. However they are of little use in isolating ribosome's and microsomes etc. These are often refrigerated. High speed centrifuges are used in more sophisticated biochemical application. Three types of rotors are available for high speed centrifugation-fixed angle, swinging bucket and vertical rotors.
- **Ultracentrifuges** - It can run at speeds up to 75,000 rpm, sufficient to allow fractionation of bio molecules, for example: plasmid DNA, chromosomal DNA and RNA. It provides 5, 00,000g centrifugal force. The Ultracentrifuge is capable of reaching even greater velocities and requires a vacuum to reduce friction and heating of the rotor. The drive shaft on which rotor is mounted is merely $1/18^{th}$ of an inch in diameter. Shaft is made up of aluminium or titanium to withstand a great force. Ultracentrifuges are usually refrigerated and are very expensive. It is used for both preparative work and analytical work. The ultra centrifuge can be classified into four types *viz*: Differential centrifugation, Density gradient centrifugation, Zonal centrifugation and Isopycnic centrifugation. In differential centrifugation separation is carried out in a homogenous

suspending medium. This is used to isolate intracellular organelles. However the disadvantage is that the yield is poor and the lighter particle near the heavier particle at the bottom also sediment resulting in contamination.

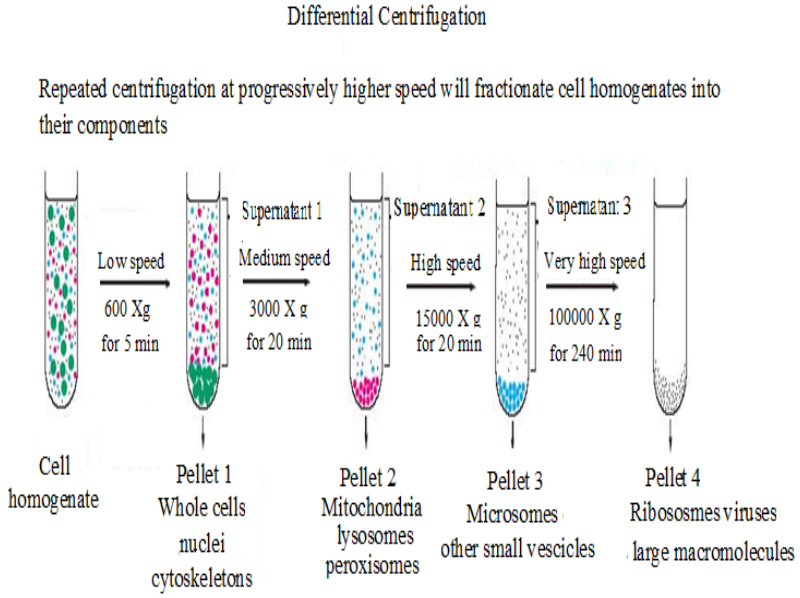


Fig: 1. Diagrammatic illustration of Differential Centrifugation

In case of Density gradient centrifugation it employs a medium which has gradient, generally a sucrose density gradient, which is created by gently overlaying lower concentrations of sucrose on higher concentrations in centrifuge tubes. The seperation under centrifugal field is therefore dependent upon the buyont densities of the particle. The particles travel through the gradient until they reach a point at which their density matches with the density of surrounding sucrose. Then the fraction is removed and analyzed. Gradient exert their seperating effect on particle and also eliminate mixing of seperated components due to convention and mechanical vibrations. There are two types in Density gradient centrifuge *viz*: Rate zonal (fig:2a and 2b) and Isopycnic. In rate zonal centrifugation

which is also known as band or gradient centrifugation, a density gradient is created in a test tube with sucrose having high density at the bottom. The gradient used here has maximum density below that of least sedimenting particle.The sample particle travel throught the steep gradient and form descreate zones depending upon their sedimentation rate, difference in size and shape.This method is usefull for seperating particle which differ in size but not in density.This is used for seperation of RNA and DNA hybrids and ribosomal subunits. In Isopycnic centrifugation the gradient used has greater densitiy than the most dense sedimenting particles.Here centrifugation is done for prolonged periods at high speed to permit all particle to seek this equilibrium density.It depend solely on the buyont densities of the particle to be seperated and not on their shape or size. Usually caesium chloride Dicoll (high molecular weight sucross polymer and epichlorohydroxy ludox (silica sols) is used as a density gradient material.

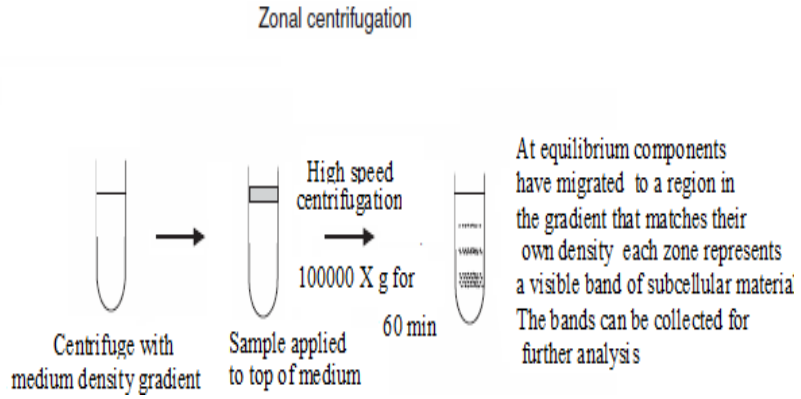


Fig: 2a. Diagrammatic illustration of Rate zonal centrifugation

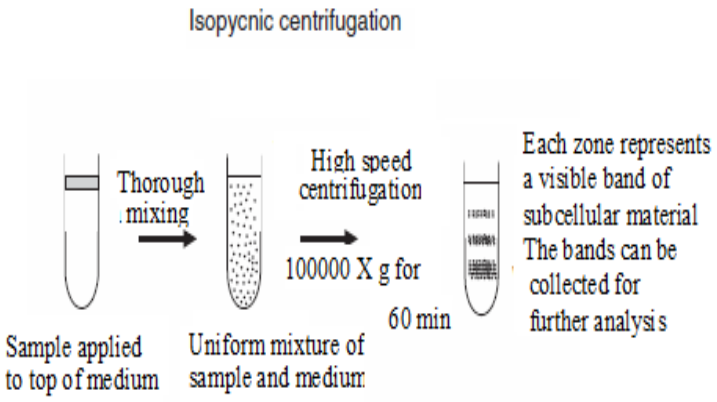

Isopycnic centrifugation
Thorough mixing
High speed centrifugation
100000 X g for 60 min
Each zone represents a visible band of subcellular material The bands can be collected for further analysis
Sample applied to top of medium
Uniform mixture of sample and medium

Fig: 2b. Diagrammatic illustration of Isopycnic centrifugation

Generally the centrifuge is used in clinical biochemistry, in production of bulk drugs, production of biological product for e.g., bacterial enzymes are separated from bacterial culture medium by sedimenting the bacterial cells by use of centrifugation. Dirt water is separated from olive and fish liver oils. Evaluation of suspensions and emulsions and determination of molecular weight of colloids.

2. Chromatography

Chromatography serves as a means of separation of mixtures and for the isolation and partial description of the components whose presence may be known or to be identified. Chromatography was first employed by the Russian scientist Mikhail Tsvet in 1900. He separated chlorophyll II in 1906 by using thin layer chromatography. Like that of centrifugation Chromatography is also of preparative and analytical type. Preparative chromatography is used to separate the components of a mixture for further studies or for purification. Whereas, analytical chromatography is done normally with smaller amounts of material and is for measuring the relative proportions of analytes in a mixture. The concept of partition coefficient is the basic principle of chromatography. In chromatography, the components to be separated are distributed between a stationary phase and the

mobile phase in which a stationary phase could be paper (paper chromatography), silica (thin layer and liquid chromatography), liquid silicone (gas liquid chromatography) etc and the mobile phase could be polar and non polar solvents and inert gases etc. Separation of the components in a sample is based on the fact that the rate of travel of an individual solute molecule through a column or thin layer of adsorbent is directly related to the partition of that molecule between the mobile and stationary phase. The partition coefficient of each component determines how much of it is in each phase at any time and how long it remains in that phase. After some time interval they will be distributed in space over the stationary phase and subsequently emerge out of the stationary phase as a single components.

Table: 2. Chromatographic techniques with mobile and stationary phase.

Mobile phase	Stationary phase	Technique
Liquid	Liquid	Partition chromatography
Gas	Liquid	Gas chromatography
Liquid	Ion exchange resin	Ion exchange chromatography
Liquid	Molecular sieves	Gel permeation, ion exclusive
Liquid	Thin layer of silica or alumina	Thin layer chromatography
Liquid	Paper	Paper chromatography

The choice of chromatographic technique solely depend on the purpose of the experiment in which the solute particle or the analytes are to be isolated, separated, estimated, to be identified, for characterization and structural elucidation etc. Hence chromate graphic techniques of various types *viz*; Column chromate graphy, Gas chromatography, High performance liquid chromatography, Affinity chromatography, Ion exchange chromatography, Thin layer chromatography, Gel permeation chromatography and etc does exist.

Table: 3. Principle and applications of chromatography

Chromatography	Principle	Applications
Column	The adsorption equilibrium or distribution coefficient between stationary solid phase and mobile liquid phase	Separation of amino acids, carbohydrates, neutral lipids, phospholipoids etc.
Gas liquid	Partition equilibrium of sample between stationary liquid phase and a mobile gas phase	To study chemical reaction rates Used to study molecular weight, bond angle deformation, ionization potential and electron affinity.
High Performance Liquid Chromatography	Partition equilibrium of sample between stationary liquid phase and a mobile liquid phase. Seperation based on mode of the sample with different mode of HPLC i.e., adsorbent, ion exchange, affinity etc.	Separation of proteins, nucleic acids, polysaccharides, plant pigment, amino acids, pesticides, steroids, drugs and their metabolites, animal and plant harmones and complex liquids.
Affinity	The equilibrium between	Used to purify the large variety of

	stationary immobilized ligand and a mobile liquid phase. It employs the capacity of biomolecules for specific non covalent binding of other molecules called ligands.	macromolecules, cells, immunoglobulin's, membrane receptors, nucleic acids and even polysaccharides.
Ion exchange	Anion exchange equilibrium between stationary ion exchange and a mobile electrolyte phase, It relies on the attraction between oppositely charged particles or the reversible exchange of ions in the solution.	Used to analyse the amino acids or called as amino acid analyser Used to determine the base composition of nucleic acids Used for exchanging magnesium, calcium, and iron ions from water and effective method for water purification Used to separate many vitamins, biological amines, organic acids and bases etc.
Thin layer	The principle of the distribution process may be based in that of adsorption, partition, chiral, ion exchange or molecular exclusion	Effectively used to identify drugs, contaminants and adulterants. Widely used to resolve plant extract and many other biochemical preparations.

	chromatography.	
Gel permeation chromatography	The equilibrium between a liquid stationary phase trapped inside the pores of a stationary porus structures and a mobile liquid phase. Separation of molecules is based on their molecular size and shape.	For the molecular weight determination of macromolecules.

Since High performance Liquid chromatography and Gas chromatography are extensively used in research and chemical industries the brief description of them is mentioned in the following session.

High pressure liquid chromatography (HPLC)

Generally the instrumentation part consists of high pressure pumps which can pump the mobile phase at 0.01 to 10ml/min at the pressure up to 500bars. HPLC contain 6 port injection valve from which sample mixture is injected. The samples should be miscible with the mobile phase. However before injecting the samples they should be filtered and processed. The mobile phase reservoirs are placed on top of the instrument. In normal phase HPLC a mobile phase is non polar e.g., hydrocarbons whereas a stationary phase is polar for e.g., AL_2O_3, $SI0_2$. In case of reverse phase HPLC mobile phase is polar such as water, acetonitrile or mixture of them and etc, and the stationary phase is non polar. The run time is usually 5 to 60 min depending on the sample. After run is completed the samples enter to the detectors. There are different types of detectors are available based on requirement of studies. UV visible or Diode array detector is used to study the larger organic molecules and transition metals which absorb the UV visible light. Fluorescence detectors are used for highly condensed organic molecules and it detects fluorescence radiation emitted by the sample compounds. Other types of detectors are Refractive index detector which is low cost and non specific, electric conductivity detector for compound dissociated into ions and Mass Selective detector.

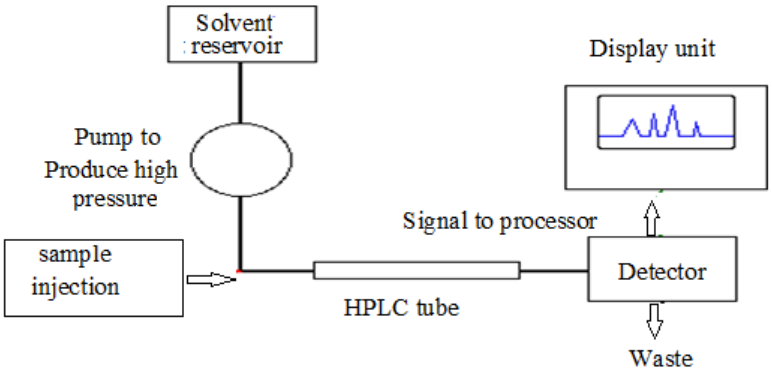


Fig: 3. Diagrammatic illustration of HPLC instrumentation.

Gas chromatography (GC)

Gas chromatography consists of flow controllers which control the flow of the mobile phase which is an inert gas such as Helium, nitrogen, argon and etc of purity 99.999%. The instrument has an injection port through which samples are injected. The samples may be in gas form or in liquid form.Theinjection port is maintained at a higher temperature than the boiling point of the least volatilecomponent in the sample mixture. However the sample which has to be analysed should be stable under GC operation conditions and should have a vapour pressure higher than zero. In standard GC the analysis or run time may be between 5 to 60 min whereas it is less than 2 min in micro GC. Once the sample injected they will be vaporized and pushed through the column and get separated. Here the columns are of two types *viz;* capillary and packed columns. Capillary columns diameter ranges from 0.1 to 0.53 mm and length from 10 to 50 meters and are made up of thin fused silica. Inside the column a thin layer of stationary phase is coated measuring about 0.3 to 50 micrometer thickness. The stationary phase could be silicon polymers (polysiloxanes Si-O-R) where R could be methyl, phenyl, cyanopropyl, ethylene glycol or fluorinated hydrocarbon. The packed columns are having a diameter of 0.53 mm

and length of 1 to 5 meters; typically they are made up of a glass or stainless steel coil. The small particles inside the column acts as stationary phase and the particle diameter would be 45-120 mesh. These small particles are immobilized in the wall for separation of high volatile compounds. Some of the stationary phase is AL_2O_3, polystyrene divinylbenzene (DVB). Once the samples get separated they will enter into the detectors. The important detectors of GC are Thermal Conductivity Detector (TCD) which is a non selective one and which can detect anything that differs from the carrier gas. Flame ionization detector (FID) Fig: 4a is far the most widely used detector. It measures all organic compounds and it can detect as low as one nanogram of any given compounds. It consists of a hydrogen/air flame and a collector plate.The efflucent from the GC column passes through the flame, which breaks down organic molecule and produces ions. The ions are collected on a biased electrode and produce an electric signal. It is extremely sensitive and has large dynamic range.Another type of detector is Electron capture detector (ECD) which has a radioactive source which ionizes the carrier gas coming out of the column. When an organic molecule that contains electornegative functional group, such as halogens, phosphorous and nitro groups, pass by the detector they capture some of the electrons and reduce the current. It is mostly used to measure polyhalogenated compounds particularly pesticides. It is very sensitive and can detect as little as one picogram of these compounds. The important and advance detector is Mass Selective Detector (MSD) Fig: 4b which is tunable for any species of sample or ions. It uses the difference in mass-to-charge ratio (m/z) of ionized atoms or molecules to separate them from each other.Itcreate gas-phase ions and separate the ions in space or time based on their mass to charge ratio then measure the quantity of ions.

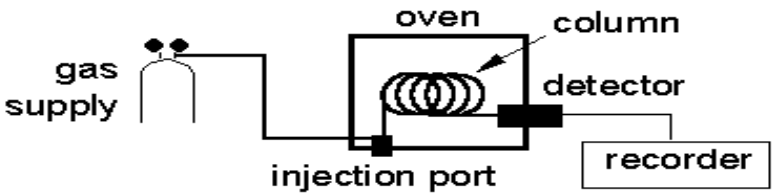


Fig: 4. Schematic representation of gas chromatography

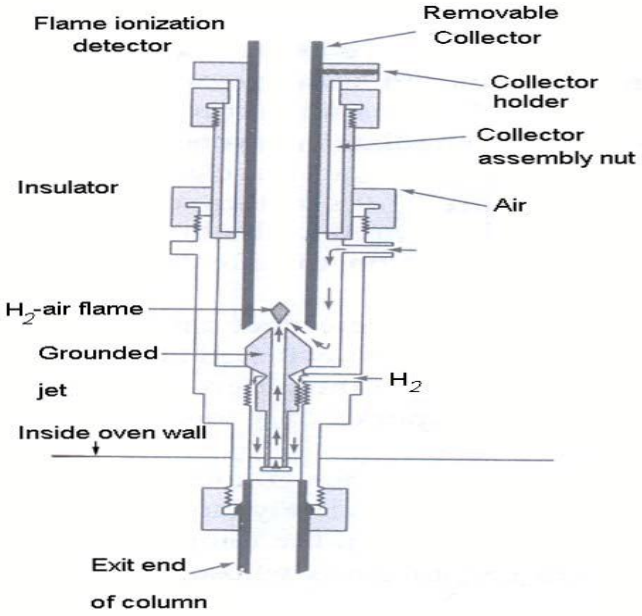


Fig: 4a. Flame ionization detector

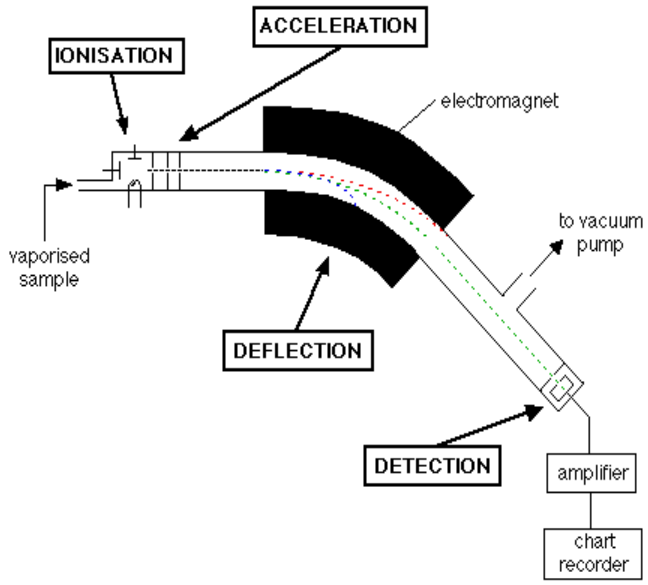


Fig: 4b. Separation of cations in Mass selective detector

Applications

Identification and quantitation of volatile and semivolatile organic compounds in complex mixtures. Determination of molecular weights and (sometimes) elemental compositions of unknown organic compounds in complex mixtures. Structural determination of unknown organic compounds in complex mixtures both by matching their spectra with reference spectra and by a priori spectral interpretation.

3. Electrophoresis

The molecules are separated on the basis of their charge and mass ratio in an electric field or in another words migration of charged ions in an electric field, this is the principle of electrophoresis. Several types of medium are present to perform the electrophoresis procedure *via*; Filter paper which is moistened with buffer and placed between electrodes and by which small molecules, amino acids and nucleotides can be separated. Polyacrylamide gel

which are covalently cross linked and are casted in tubes or as slabs, they are used to separate proteins and nucleic acids. Agarose gel a highly purified polysaccharide derived from agar (extracted from seeweed), long sugar polymers held together by hydrogen and hydrophobic bonds but with no crosslinks are used to separate very large proteins, nucleic acids and nucleoproteins. Two types of gel electrophoresis include; one dimension and two dimension. One dimension includes SDS-PAGE, Native PAGE and Iso Electric Focussing (IEF). SDS-PAGE, the most widely used electrophoresis technique, separates proteins primarily by mass. Non denaturing PAGE, also called native PAGE, separates proteins according to their mass/charge ratio. The two-dimensional PAGE (2D-PAGE) separates proteins by isoelectric point in the first dimension and by mass in the second dimension.

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).

SDS-PAGE is a method of gel electrophoresis to separate proteins based on the mass. Sodium dodecyl sulphate (SDS) is an anionic detergent that breaks up the interactions between proteins. This also disrupts secondary and tertiary protein structures by breaking hydrogen bonds and unfolding protein. SDS binds to protein in a ratio of one SDS molecule/two amino acids. SDS will give a total negative charge to all the protein molecules irrespective of their charges. The negatively charged proteins denatured by SDS are applied to one end of a layer of polyacrylamide gel submerged in a buffer. The buffer provides uniform pH and ions for conducting electric potential. Some of the common buffers used are Barbital buffer & Tris-EDTA, Tris-acetate-EDTA and Tris-borate-EDTA. When an electric current is applied across the gel, the negatively charged proteins migrate across the gel to the positive pole. The gel used for SDS-PAGE is made out of acrylamide which form cross-linked polymers of polyacrylamide. Standard gels are typically composed of

two layers, one top most layer called the stacking gel and a lower layer called separating or resolving gel. The stacking layer contains a low percentage of acylamide and has low pH, while the acrylamide concentration of the separating gel varies according to the samples to be run and has higher pH. The difference in pH and acrylamide concentration at the stacking and separating gel provides better resolution and sharper bands in the separating gel. Short proteins will more easily fit through the pores in the gel and move fast, while larger ones will have more difficulty. Due to differential migration based on their size, smaller proteins move farther down the gel, while larger ones stay closer to the point of origin. After a given period of time, proteins might have separated roughly according to their sizes. Standard proteins were also run along the side of the test sample for molecular weight determination. Following electrophoresis, the gel may be stained with Coomassie Brilliant Blue or silver stain to visualize the separated proteins. After staining, different proteins will appear as distinct bands within the gel according to their sizes (and therefore by molecular weights).The molecular weight of a protein in the band can be estimated by comparing it with the marker proteins of known molecular weights. The relative mobility of each protein band is calculated as,

$$R_m = \frac{\text{distance protein moves}}{\text{distance dye moves}}$$

Native PAGE is used to separate proteins in their native states according to difference in their charge density. No denaturants will be added to the gel and buffer. It is used for preparation of purified and active proteins. In native PAGE the mobility depends on both the protein's charge and its hydrodynamic size. The charge depends on the amino acid composition of the protein as well as post- translational modifications. Iso electric Focusing (IEF) is a technique for separating different molecules by

their electric charge differences. In this technique proteins are separated by electrophoresis in a pH gradient based on their isoelectric point (pI). A pH gradient is generated in the gel and an electric potential is applied across the gel. At all pHs other than their isoelectric point, proteins will be charged. If they are positively charged, they will move towards the more negative end of the gel and if they are negatively charged they will move towards the more positive end of the gel. At its isoelectric point, since the protein molecule carries no net charge it focuses into a sharp band. In two dimensional electrophoresis the first step is to perform IEF and then SDS-PAGE. Proteins having same molecular weight or same pI are also resolved. This technique can be used for analysis of complex mixture of proteins, partial characterization of proteins. Use of the electrophoresis include, to determine molecular weight of a peptide/protein, to identify protein, to determine sample purity, to identify existence of disulfide bonds, to quantify amounts of protein

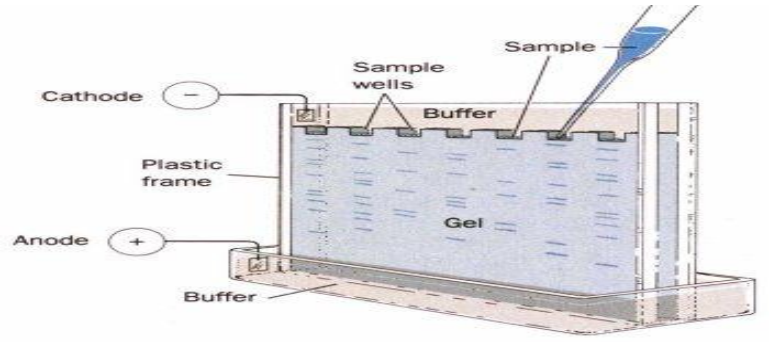


Fig:5.Diagram of polyacrylamide gel electrophoresis

One of the best uses of electrophoresis is its use in DNA and protein fingerprinting *viz;* Southern, Northern and Western blotting. Western blot (also called as immunoblot) is a technique to detect specifically one protein in a mixture of large number of proteins and to obtain information about the size and relative amounts of the protein present in different samples. At first, proteins are separated using SDS-polyacrylamide gel electrophoresis. Then they are moved

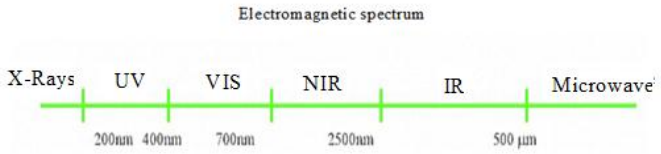
onto a nitrocellulose membrane. The proteins retain the same pattern of separation they had on the gel.An antibody is then added to the solution which is able to bind to its specific protein and forms an antibody-protein complex with the protein of interest. Finally the nitrocellulose membrane is incubated with a secondary antibody, which is an antibody-enzyme conjugate that is directed against the primary antibody. The location of the antibody is revealed by incubating it with a substrate that the attached enzyme converts to a product that can be seen and followed and then photographed. In southern blotting genes can be localized by means of hybridization to radioactive DNA or RNA molecule (probe) which has a complementary sequence. Southern blotting was developed by Edwin Southern in 1975 it is also called as DNA blotting or hybridization. It is especially used to analyze the DNA or sometimes also called as DNA fingerprinting. In its first step the DNA from the source is extracted and cut into many pieces by using restriction enzymes. The segmented DNA are run in an agarose gel electrophoresis and stained with ehidium bromide. The gel is taken out and placed on neutralized nitrocellulose paper or nylon filter. Then to this 32p radio labeled complimentary DNA or RNA in a minimum solution volume at 68 ^{o}c of higher ionic strength is added which hybridizes with DNA, this is called hybridization. Then the paper is covered with X ray film or photographic plate which, when developed, reveals the position of the radioactive probe. The difference between northern and southern blotting is, in case of northern blotting which is used for the study of RNA, alkali is not used since the alkali hydrolyses the RNA instead formaldehyde is used. Since RNA doesn't bind to nitrocellulose paper unless it is denatured so DBM or **diazobenzyloxymethyl-pape**r is used. Northern blotting was developed by James Alwine, David Kemp and George Stark in 1977. With this technique quality and quantity of gene can be measured.

4. Spectroscopy

Atoms and molecules interact with electromagnetic radiation and may absorb and/or emit Electromagnetic radiation.The patterns of absorption and/or emissions are called '*spectra*'. Spectroscopy is concerned with the interpretation of these spectra. In other words it is a study of interaction between electromagnetic radiation and matter. Different regions of the electromagnetic spectrum provide different kinds of information as a result of such interactions. A spectrophotometer is an instrument that measures the amount of light or electromagnetic radiation that is absorbed or emitted by a sample. Arnold J. Beckman at the National Technologies Laboratory (NTL) invented the Beckman DU spectrophotometer in 1940. Some of the electromagnetic wave parameters is useful in understanding the spectroscopy.Wavelength (λ): Wavelength is the distance between the consecutive peaks or crests and is expressed in nanometers 1nm=10^{-8} meters.Frequency (ν): Frequency is the number of waves passing through any point per second and is usually expressed as Hertz(Hz).Wave number (ν): Wave number is the number of waves per cm.Wavelength, Wave number and Frequency are interrelated as,

$$1/\lambda = \bar{\nu} = \nu/c$$

Where, λ is wave length, $\bar{\nu}$ is wave number, ν is frequency, c is velocity of light in vacuum. i.e., 3 x 10^{-8} m/s



The principle of the spectroscopy is based on Beer–Lambert law / Beer –Lambert – Bouguer law which states that the amount of light absorbed is proportional to the concentration of the absorbing substances and to the thickness of the absorbing material (path length).

$$\log (I_0/I) = \varepsilon c l = A$$

Where, I_0- the intensity of incident light I- the intensity of transmitted light, ε - molar absorptivity / molar extinction coefficient in cm 2 mol^{-1} or L mol^{-1} cm^{-1}. c - concentration in mol L^{-1}, l - path length in cm, A- absorbance (unitless).

This law can be used to find the concentration of solutions absorbing in UV or visible region. However there are deviations from this law or the limitations of the Beer–Lambert law is that Beer's Law successfully describes the behaviour of dilute solutions only. At high concentrations (i,e. greater than 10^{-2} M) there is interaction between absorbing particles such that the absorption characteristics of the analyte are affected. The voltage fluctuation, sensitivity changes in the detector and solution containing more than one complex which absorb different wavelength are the other deviating factors of this law. Spectroscopy can be used for qualitative analysis and quantitative analysis. In qualitative analysis the characteristic wavelength are used for a given analyte and through which we can identify the sample and such type of spectrometers are called as photometers (e.g., HPLC detectors). Whereas in quantitative analysis the intensity of absorption or emission from the analyte is used to find out the concentration of the analyte in a given sample and quantitative measurements can be made at any desired wavelength, such type of spectrometers are called as spectrophotometer (e.g.,UV-Visible spectrometer). There are three types of spectrophotometer *viz;* single, double and split beam spectrophotometer. In single beam all the light passes through the sample and to measure the intensity of the incident light sample

must be removed so that all the light can pass through. It is cheap and less complicated. In double beam spectrophotometer it compares the light intensity between two light paths, one path containing a reference sample and the other the test sample. The readings are stable. The disadvantagesare higher cost, lower sensitivity because of the more complex optics. Split beam is similar to the double beam spectrophotometer but it uses a beam splitter instead of a chopper to send light along the blank. The essential components of a typical spectrometer are,

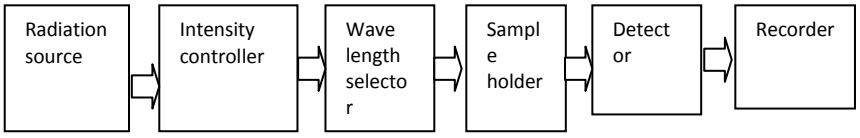


Table:4.Componenets of spectrometer

Radiation source	Dispersion device	detector	Used in
Tungsten lamps,	Monochromators Prism or diffraction gratings +slit Filters etc	Phototubes Photo multiplier Diode array Thermo couple Pyroelectric Photo electric cells	Visible spectrophotometer Colorimeters and HPLC detectors
Deuterium, mercury or hydrogen lamps			UV spectrophotometer
Mercury,Xenon flash lamps,Deuterium, tungsten.			UV-Visible spectrometer
1.1.2 X ray tube**,** Synchrotron Radiation Source (SRS), Betatron (cyclotron)			X ray spectroscopy
Flame, plasma, hollow cathode lamp			Atomic emission and absorption spectrophotometer
Electrically heated rod of rare			Infra Red spectrophotometer

earth oxides, Silicon carbide globar			
Magnet of stable field strength			Nuclear Magnetic Resonance
Gamma-emitting nuclides			Gamma spectrometer
Klystron valve			Electron Spin Resonance spectroscopy

Some of the common types of analytical spectroscopes are; Absorption, Fluoresence and Phosphorescence, Emission (atomic with flames, arcs, sparks, and plasmas), Chemilumenesence and Bioluminescence and Reflection spectroscopy. Few important spectroscopy methods are described here.

UV-Visible spectroscopy

The electronic spectra of organic molecules which are found in UV region (100nm-400nm) and visible region (400nm-750nm) can be studied by the UV-Visible spectrophotometer. UV and Visible radiations absorbed by the molecules will bring transition of outer shell electrons. When the sample is irradiated with the UV - Visible radiation if a particular electronic transition matches the energy of a certain band of UV - Visible, it will be absorbed. The remaining UV - Visible light passes through the sample and is observed with the gaps in the spectrum region and is called absorption spectrum.

Applications

- To determine the absorbance or transmission of characteristic wavelengths of radiant energy (light) by a chemical species in solution.
- Identify organic compounds by determining the absorption maximum.
- Used for color determination within the spectral range
- Concentration of impurities
- To study the rate of a reaction

Infra Red Spectropscopy (IR spectroscopy)

Infra red has long wavelength and less energy than UV and visible spectrum. It is associated with vibrational transition but not with the electronic transition. Infrared radiation stimulates molecular

vibrations and absorption of radiation by a sample is due to changes in the vibration energy states of a molecule. Frequency of the incident radiation is varied and the quantity of radiation absorbed or transmitted by the sample is obtained.The spectra are represented as percent transmittance versus wave number.

Applications

- It is useful in identifying functional groups of a compound.
- Stretching of bonds, bending of bonds, or internal rotation around single bonds can be revealed.
- The measurement of the type and amount of light transmitted by the sample gives information about the structure of the molecules comprising the sample.

Fluorometry

Flurometry is based on the phenomenon whereby a molecule after absorbing radiation emits radiation of longer wavelength which is known as fluorescence. It is a short lived phenomenon about 10^{-7} seconds. Fluorometry is an analytical tool which can be used for determination of very small concentration of substances. As that of UV Visible spectroscopy Beer Lamberts law is also applied to the Fluorometry. Some of the applications of spectrofluorometry are;

- Qualitative analysis
- Quantitative analysis of vitamins, cortisol, serotonin, dopamine etc
- Assay of oraganophosphorus pesticides, carcinogens, drugs and even some metal ions.

Flame spectroscopy

It is the absorption or emission of specific wavelength by exited atoms which is taken by flame. There are two types in flame

spectroscope *viz*; Emission Flame photometry and Atomic Absorption Spectrophotometry. In Emission Flame photometry, the characteristic emission spectrum of the element is produced when the excited atoms return to their ground state. Whereas, in Atomic Absorption Spectrophotometer it measures the absorption of a beam of monochromatic light by atoms in a flame.

Applications

- Assay of macro and microelements in blood, plasma, urine, saliva, CSF, milk, tissues, soil samples and plants.
- Flame photometry is especially used in estimation of alkali, alkaline earth and rare earth materials.
- They can be used to estimate gold, silver, iron, lead, copper, zinc and other elements in the biological samples and as well as in animal feed and plant materials

Mass Spectrometry

It uses high energy electrons to break the molecule into fragments or ions either positively or negatively charged or commonly a cation and a radical. Bonds break to give the most stable cation. Stability of the radical is less important. Only cations are detected where as radicals are "invisible" in MS. These charged particles can be manipulated in an electric or magnetic field depending on their *mass* (m) and the *charge* (z) of the particle. It operates under high vacuum (keeps ions from bumping into gas molecules).Most cations formed have a charge of +1 so the amount of deflection observed is usually dependent on the mass of the ion. The resulting mass spectrum is a graph of the mass of each cation vs. its relative abundance.

Applications

- Molecular mass, weights and molecular structure (fragmentation) determinations.

- Discovery of isotopes and characterization of new elements
- Qualitative and quantitative analyses
- Identification of reaction products and industrial products for quality control
- Identification of trace elements, pollutants in drinking and wastewater and drugs
- Useful in the identification of a newly synthesized compound.

The ion spectrum of unknown compound can be compared with the absorption spectrum of several known compounds in the MS library and the unknown compound correlates or matches with the known compound then further confirmation of the compound cab be done by use of FTIR and NMR spectroscopy. The matching of the spectra with the standards of known compound in the MS library is called as spectral fingerprinting.

Nuclear Magnetic Resonance (NMR) spectroscopy

In NMR spectroscopy the phenomenon of Nuclear magnetic resonance is employed for the study of physical, chemical and biological properties of the analyse. Atomic nuclei possess spin and because which it generates a magnetic field and gains the magnetic moment. The frequency of spinning nucleus is exactly equal to the frequency of Electromagnetic radiation necessary to induce a transition from one nuclear spin state to another. When these magnetically active nuclei are placed into an external static magnetic field, the magnetic fields align themselves with the external field into two orientations. When electromagnetic radiation of specific frequency is applied, by sweeping the magnetic field, an energy difference between spin states will occur that has the same energy as that of the applied radio frequency and plot of frequency versus energy absorption can be generated. The nuclei will

absorb radiation in the radio wave region of the electromagnetic spectrum, this is the NMR spectrum.

Applications

- To study the structure of small organic molecules and small globular proteins.
- Used for 3-Dimensional structure determination of macromolecules.
- Chemical shift mapping – Structural information on the binding modes and site positions.
- Molecular dynamics, conformational analysis.
- Use of 3D-NMR &single-slice planar in medical field such as imaging in detecting breast abnormalities etc.

Summary

All biomolecules or macromolecules have definite functions by virtue of their structural and chemical properties in the biological or living system. They contribute the major component of the cell. There are four important macromolecules *viz*; carbohydrates, lipids, protein and nucleic acids. These are basic requirements for the metabolic and catabolic pathways of the cell. They are involved in energy process, signalling, transportation, receptors, defence, carrying genetic information and various other cell functions which are essential for life. The lack of these macromolecules or any undesired changes in the chemical or structural properties of these macromolecules will lead to abnormalities, mutation and can influence adverse affects on the living system or life.

The analytical tools or methods have gained the importance in modern biological experiments since they can be exclusively employed for the study of biomolecules. They are helpful in assessing the biological conditions of the living system. These tools are extensively used in

medical, pharmaceutical and research fields. Not only Centrifugation, Spectroscopy, Chromatographic technique and Electrophoresis are important in analytical tools, but when they are paired with other kinds of analytical tools they prove to be the versatile techniques to study the various aspects of macromolecules. These techniques can reveal the structural, functional, chemical, physical and biological functions of the macromolecules. By the study of macromolecules with the aid of different analytical tools the modern science has gained the access to make suitable or appropriate changes in the living system as and when required.

Suggested readings

1. David Freifelder molecular biology 2nd edition Jones and Bartlett publishers USA
2. D. J.Taylor,NPO Green and GW Stout Biological science R Soper (Ed) 1997 Cambride University press
3. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T. Nature 277, 680–685.
4. Lehninger Principles of Biochemistry by David L Nelson and Michael M cox Third edition 2000 Worth publisher New York
5. Mikhail Tswett (1906) "Physikalisch-Chemische Studien über das Chlorophyll. Die Adsorption."(Physical-chemical studies of chlorophyll. Adsorption.) *Berichte der Deutschen botanischen Gesellschaft*, vol. 24, pp. 316–326.
6. Modern and experimental biochemistry-RODNEY BOYER
7. Principles and technique in biochemistry-L WALKER& WILSON
8. Upadhyay, Upadhyay and Nath Biophysical Chemistry-Principle and Techniques. Himalaya publishing House.

Chapter 4
Plant Molecular Farming: A Promising Stratergy in Biotechnology

Harunipriya. P, Chakravarthi. M, Brindha. C
and Chandrashekara. K. N

The improvement of agricultural crop plants relied largely on the conventional breeding programs to increase the productivity, alter the quality characteristics or impart resistance to biotic or abiotic stresses. However, with the advancement of molecular biology techniques it has become possible to introduce entirely new characteristics efficiently through insertion of the genes coding for the desired characteristics directly into the genome of the plant or animal. Also, these techniques have made possible to introduce non-native genes from varied sources, thus tailoring the plant or animal to produce entirely new products for innovative applications. A large number of recombinant proteins have been produced in plants over the last twenty years, demonstrating the ability of plants to compete with existing industrial production systems. The use of plants for producing recombinant proteins has been termed as “Molecular Farming” and its rapid progress is driven by the several advantages provided by the plant expression systems. In this review we discuss the advantages of using plant as expression systems, type of biomolecules and recent examples, the plant endomembrane system, downstream processing of recombinant proteins in plants, different novel plant-derived pharmaceuticals and non-pharmaceutical protein

products that are at various stages of clinical development or commercialization thereby making plant based production system suitable alternatives to the existing system. It also attempts to overview the challenges and future of plant molecular farming technology and biosafety of molecular farming products which are crucial to eliminate the potential health and environmental risks, thus ensuring the success of this system.

Introduction

Plants as Expression Systems

Plants are modified to produce a wide range of heterologous proteins including pharmaceutical and industrial proteins, through recombinant DNA technology, often referred as plant molecular farming (Faye and Gomord, 2010; Ma and Wang, 2012; Obembe *et al*., 2011; Wilken and Nikolov, 2012). As green bioreactors, plants offers a variety of advantages such as nearly unlimited scalability, from small scale trials in growth chambers to large open-field mass production, and all at relatively inexpensive cost. Plants have become a promising alternative over the traditional expression systems to produce a variety of valuable biological molecules ranging from medicinal applications such as vaccines to materials like biodegradable plastics with industrial uses (Twyman *et al.,* 2005).Plants can produce sufficiently high yields of proteins than bacterial or yeast fermentation systems and at 0.1% of the cost of mammalian cell cultures (Twyman *et al.,* 2003). In addition plants have an advantage over other protein expression systems, such as bacteria, for the production of antibodies and other complex proteins because they are able to make, fold and correctly assemble proteins consisting of multiple subunits. As an example, secretory Immunoglobulin A (sIgA) which consists of four linked proteins are successfully produced in tobacco plants (Goldstein and Thomas, 2004). The comparison of recombinant protein production in plants, yeast and mammalian systems (Ma *et al*., 2003) is given in table 1.

Table 1. Comparison of recombinant protein productions in different systems

System	Cost	Production Time	Scale up Capacity	Product Quality	Glycosylation	Contamination Risks	Storage Cost
Bacteria	Low	Short	High	Low	None	Endotoxins	Medium
Yeast	Medium	Medium	High	Medium	Incorrect	Low risk	Medium
Mammalian cell culture	High	Long	Very low	Very high	Correct	Virus, prions, oncogenic DNA	High
Transgenic animals	High	Very long	Low	Very high	Correct	Virus, prions, oncogenic DNA	High
Plant cell culture	Medium	Medium	Medium	High	Minor	Low risk	Medium
Transgenic plants	Low	Long	Very high	High	Minor	Low risk	Very low

Types of Biomolecules and recent examples

In recent years, several important products such as human biopharmaceuticals, recombinant antibodies, recombinant subunit vaccines, nutritional supplements, biodegradable plastics have been produced in plants with high success (Miao *et al*., 2008). The first pharmaceutically relevant protein made in plants was human growth hormone, which was expressed in transgenic tobacco in 1986 (Barta *et al*., 1986). The first antibody was also expressed in tobacco in 1989, which proved that plants could assemble complex functional glycoproteins with several subunits (Hiatt *et al*., 1989). Since then, other important vaccine candidates and therapeutic proteins have been produced in transgenic plants and are in different stages of clinical trials (Ma *et al*., 2003). Some important recombinant products produced in plant systems is given in table 2.

Table 2. Transgenic plant based products available in market

Product	Plant system	Company name	Commercial name
Aprotinin	Corn, tobacco	Prodigene	AproliZean
β-glucuronidase	Corn	Prodigene	GUS
Trypsin	Corn	Prodigene	TrypZean™
Recombinant human lactoferrin	Corn, rice	Meristem Therapeutics	Lacromin™
Recombinant human lysozyme	Rice	Ventrica Biosciences	Lysobac™
Recombinant lipase	Corn	Meristem Therapeutics	Merispase™

Avidin	Corn	Prodigene	Avidin
Recombinant Human intrinsic factor	*Arabidopsis*	Cobento Biotech AS	Coban
Collagen	Corn	Prodigene, Medicago	---

Plants can be engineered to act as bioreactors for vaccines and therapeutic production, and their metabolic pathways can be manipulated to increase compounds of benefit or decrease detrimental compounds. Plant molecular farming applications can be broadly classified as plant made pharmaceuticals (PMPs), Plant made nutritional compounds (PMNs) Plant made vaccines (PMVs), Plant made plastics (PMPs) and plant made industrials (PMIs) (Horn *et al*., 2004).

Plant made pharmaceuticals (PMPs)

Therapeutic proteins are bioactive molecules that have potential applications in medicinal diagnostics and therapy. Several therapeutic products can be produced in plants which include diagnostic proteins (antibodies and enzymes), replacement proteins (Factor VIII for hemo philiacs, in sulin for diabetics), immune system stimulator/suppressants (interleukins, interferons, and colony stimulating factors), and adhesive proteins for surgical purposes or for growth factors (Daniell *et al.,* 2001b; Goldstein and Thomas, 2004; Rajasekharan, 2006; Twyman *et al*., 2005). Antibodies or immunoglobulins (IgGs) are serum proteins that play a central role in the humoral immune response and the production of these antibodies in plants are referred as “Plantibodies” (De Jaeger *et al.,* 2000; Goldstein and Thomas, 2004). The production of immunoglobulin fragments and their assembly in plants was reported in tobacco for the first time (Hiatt *et al*., 1989).

The recombinant antibodies can be produced in plants in many forms which include full size recombinant antibody, chimeric antibody, secretory antibody, single chain Fv fragments (scFvs), scFv fusion, bispecific scFv, antibody fragments or heavy chain variable domains (Ma *et al.,* 2003). Transgenic plants have also been used for the production of antibodies directed against Dental caries, Rheumatoid arthritis, Cholera, Diarrhoea, Malaria, certain cancers, Norwalk virus, HIV, Rhinovirus, Influenza, Hepatitis B virus, and Herpes simplex virus (Thomas *et al.,* 2002). Commercialization of plant-derived pharmaceutical proteins for the treatment of human diseases is given in table 3.

Table 3. Plant-derived pharmaceutical proteins for commercialization for the treatment of human diseases

Product Product	**Class**	**Indication**	**Company/Organization**	**Crop**	**Status**
Various single chain Fv antibody fragment	Antibody	Non-Hodgkin's lymphoma	Large Scale Biology	Tobacco	Phase I
CaroRx	Antibody	Dental caries	Planet Biotechnology Inc.	Transgenic tobacco	Phase II
E.coli heat labile toxin	Vaccine	Diarrhoea	Prodigene Inc. Arntzen group	Transgenic maize	Phase I

Gastric lipase	Therapeutic enzyme	Cystic fibrosis, pancreatitis	Meristem Therapeutics Arntzen group (Richter *et al*, 2000)	Transgenic potato Transgenic maize	Phase I Phase II
Hepatitis B Virus surface antigen	Vaccine	Hepatitis B	Thomas Jefferson University/Polish Academy of Sciences	Transgenic lettuce	Phase II
Human intrinsic factor	Dietary	Vitamin B12 deficiency	Cobento Biotech AS	Transgenic *Arabidopsis*	Phase II
Lactoferrin	Dietary	Gastrointestinal infections	Meristem Therapeutics	Transgenic maize	Phase I
Norwalk virus capsid protein	Vaccine	Norwalk virus infection	Arntzen group (Tacket *et al.*, 2000)	Transgenic potato	Phase I
Rabies glycoprotein	Vaccine	Rabies	Yusibov *et al.,* (2002)	Viral vectors in spinach	Phase I

Human glucocerebrosidase (hGC) is required for enzyme replacement therapy was produced in tobacco and in transgenic carrot cells by Protalix Biotherapeutics. Human somatotropin (hST) which is used to treat hypopituitary dwarfism in children, turner syndrome and

chronic renal failure, was produced in tobacco chloroplasts and proteins accumulated up to 7% of TSP (Staub *et al.,* 2000). An artificial substitute for human breast milk with the bioactive proteins, a synthetic human lactoferrin (HLF) was produced in rice and the transgenic rice plants showed HLF accumulation 0.5% in dehusked rice grain (Nandi *et al*., 2002). Plants have the potential to produce large amounts of mAbs, with low production costs, the ability to be rapidly scaled up to meet market demand and reduced risk of contamination with human and animal pathogens (Fischer *et al*., 2003; Teli and Timko, 2004). Other crop plants like potatoes, alfalfa and rice have also been used to produce antibodies (Goldstein and Thomas, 2004).

Plant made nutritional compounds (PMNs)

Plants can provide most of the nutrients required in the human diet. However, major crops have been found to be deficient in one or the other nutrients. The advances in genetic modifications have made it possible to enhance the nutritional quality of the plants (Galili *et al*., 2002; Zimmermann and Hurrell, 2002). Several technical advances have been made from earlier attempts to simultaneously manipulating multiple steps in plant metabolic pathways and in constructing novel, multi-enzyme pathways in plant tissues (Kinney, 2006; Sandmann *et al*., 2006; Wu *et al*., 2005). In the last few years, a lot of progress has been made in the field of biofortification, specifically palnts with higher β-carotene, lycopene, vitamins, flavonoids, resveratrol, polyamines, nutraceuticals, amino acids, nutritional proteins, minerals, fatty acids, and carbohydrates are being produced.

Plant made vaccines (PMVs)

A vaccine is an antigenic preparation used to establish immunity against a disease and the main aim of the vaccination is to eradicate infectious diseases. In the beginning (Prakash, 1996; Artnzen, 1997)

proposed that plants can be engineered for the production of vaccines. Several antigenic determinants belonging to various pathogens causing variety of diseases including bacterial and viral diarrhea, anthrax, rabies, cancer, SARS, measles, HIV, diphtheria, pertusis, tetanus, tuberculosis, respiratory syndrome, Alzheimer's disease, malaria, foot and mouth disease of cattle, gastroenteritis, hemorrhagic disease, bursal disease, goat plague, rinder pest virus, cytomegalovirus infections, parvoviral infections of dogs, avian influenza and bovine pneumonia have been produced in plants (Khandelwal *et al.,* 2003; Sharma *et al.,* 2004; Streatfield and Howard, 2003; Tiwari *et al.,* 2009; Youm *et al.,* 2008). In recent years, a large number of antigens have been produced in plants and shown to activate the immune response against the antigen in the animal models. Dow Agro Sciences has received the first ever regulatory approval for plant-made vaccine from the USDA Center for Veterinary Biologics (CVB) in January 2006.

In addition to the production of vaccines for diseases of humans in plant systems, attempts were also made to develop transgenic plants for the production of veterinary vaccines too. Attempts have been made for the production of vaccines for foot and mouth disease (FMDV), bovine rotavirus disease virus (BRV) and bovine viral diarrhoes virus (BVDV) in plants (Santos and Wigdorovitz, 2005). Even though plant based expression system is a very attractive alternative to the conventional methodologies, low expression level of the antigen in plants is the main drawback.

Plant made plastics (PMPs)

Plastics difficult to dispose off and continually accumulating non-degradable wastes have become a significant source of environmental pollution (Shimao, 2001). Biodegradable plastics seem to be a viable alternative to synthetic plastics. The biodegradable materials can

undergo decomposition into carbon dioxide, methane, water, inorganic compounds with the help of the enzymatic actions of microorganisms within a specified period of time (Anderson and Dawes, 1990). Polyhydroxyalkanoates (PHAs) are the biodegradable polymers which occur naturally in plants. Plants have been engineered to produce PHAs or PHBs in the various plant cell compartments (John and Keller, 1996; Matsumoto *et al.,* 2009). For the economic feasibility of transgenic plants-derived biodegradable plastics, accumulation of at least 15% of the tissue dry weight is required (Scheller and Conrad, 2005). The expression level of biodegradable plastic-like compounds in plants, have also been targeted to chloroplast (Bohmert *et al.,*2000; Lossl *et al.,*2005; Lossl *et al.,*2003; Nawrath *et al.,*1994) and expression levels ranging up to 40% of dry weight have been obtained (Bohmert *et al*., 2000). It was also produced in peroxisomes by using RAVAL residues encoding sequence at the carboxy terminal, accumulating PHBs up to 2% dry weight (Hahnn *et al.,* 1999). There is a need for further improvements in PHB production in plants, aimed at higher accumulation without any side effects in plants.

Plant made industrial compounds (PMIs)

Plant molecular farming is starting to become a viable new industry. This group includes hydrolases, encompassing glycosidases and proteases, milk proteins ß-casein, lactoferrin and lysozyme, protein polymers tissue replacement (Ma *et al.,* 2003). Expression of thioredoxin in foods such as cereal grains would increase the digestibility of proteins and thereby reduce their allergenicity (Thomas *et al*., 2002). Human collagen can be produced in transgenic tobacco plants and that the protein is spontaneously processed and assembled into its typical triple-helical conformation (Ma *et al.,* 2003). The production of chicken egg white avidin in transgenic corn using an avidin gene with codon

optimization was achieved (Hood *et al.,* 1997). The endoplasmic reticulum of transgenic tobacco and vacuole of potato tubers expressed recombinant dragline silk protein up to 2% of TSP (Scheller *et al.,* 2001). Several other products have been produced in plants which include streptavidin, acetyl cholinesterase, hirudin, protein C, human β casein, vegetable oils, collagen, gamma-amino butyric acid, β-glucuronidase, cyclodextrins, enzymes like phytases, xylanases, amylase, laccase, glucanases and trans glutaminases (Akama *et al.,* 2009; Boehm, 2007; Cahoon *et al.,* 2007)

For efficient production of recombinant products, the selection of the host plant plays an important role (Sharma and Sharma, 2009). Apart from this, the life cycle of the host, biomass yield, containment, scale-up costs, the form of recombinant protein, ease of downstream processing are the deciding factors. The site of protein localization in the plant cell is another important criterion which decides the correct protein folding and its yield. Various plant organs (leaves, roots, seeds) and plant cell compartments (endoplasmic reticulum, vacuole, chloroplast, oil bodies) are being tested as sites for recombinant protein accumulation (Goldstein and Thomas, 2004).

The plant endomembrane system

A plant cell contains upto 10,000 different kinds of proteins. Each of these proteins must be localized to the precise intracellular membrane, organelle or directed to the cell surface for its proper functioning. The plant endomembrane system is a complex system of organelles specialized for the synthesis, transport, modification and secretion of proteins and other macromolecules. This system is composed of several functionally distinct membrane compartments: the endoplasmic reticulum (ER), the Golgi apparatus including the trans-Golgi network (TGN), secretory vesicles, the vacuole/lysosome and

endosomes. Notably, the membranes of mitochondria and chloroplasts do not belong to the endomembrane system (Vitale and Galili, 2001; Dacks *et al*., 2009).

However, the synthesis of the majority of proteins of a eukaryotic cell occurs in the cytosol, and from there proteins are migrated to reach their final destination. These proteins thus contain the information necessary to be transported to the correct target compartment. Targeting to the cell secretory pathway, in particular, has been proposed to improve the stability and yield of several proteins (Ma *et al.,* 2003; Yoshida *et al.,* 2004; Vitale and Pedrazzini, 2005). In the absence of any specific targeting signals, a protein entering the endomembrane system will follow the default secretory pathway and will be secreted to the cell exterior.

Subcellular Targeting of proteins in plants

Organelle-specific protein targeting, protein sequestration in, or targeting to a specific cell compartments has also been readily recognized as a key factor determining the overall stability and yield of recombinant proteins in plants (Wandelt *et al.,* 1992; Schouten *et al.,* 1996; Gomord *et al.,* 1997). Targeting signals can be used to intentionally retain recombinant proteins within distinct compartments of the cell to protect them from proteolytic degradation, preserve their integrity and to increase their accumulation levels (Seon *et al.,* 2002). Several subcellular compartments have been considered as possible destinations for recombinant proteins in plant cells, endoplasmic reticulum, chloroplast and different subcompartments of the cell secretory pathway (Ma *et al.,* 2003; Daniell, 2006; Goulet and Michaud, 2006). Recombinant products and their localization sites in the plant are depicted in figure 1.

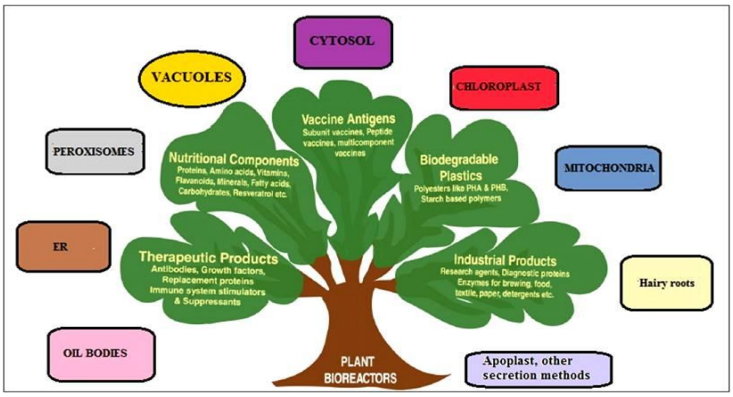


Fig 1. Recombinant products and their localization sites in the plant

Targeting to cytoplasm (modified from Sharma and Sharma, 2009)

The absence of a targeting signal in the transgene sequence prevents migration of the recombinant protein out of the cytosol following mRNA translation. Recombinant proteins retained in the cytosol are usually detected at very low levels, accumulation rates below 0.1% of TSP (Conrad and Fiedler, 1998). Several factors may explain the limited suitability of the cytosol as a destination for recombinant proteins: The negative redox potential of the cytosolic milieuis unfavorable to the correct folding of proteins with disulphide bonds (Goulet and Michaud, 2006); The important co and post-translational modification processes, such as glycosylation and the effective housekeeping activity of the ubiquitin–proteasome proteolytic pathway which may have a positive impact on the folding and assembly, structural stability of several nascent and mature proteins was absent in this compartment (Faye *et al.,* 2005, Vierstra, 1996, 2003).

Cytosolic targeting of the tomato mosaic virus antibody ‘rAb29’ in tobacco leaf cells, for instance, resulted in very weak accumulation rates, whereas the same transgene including a signal peptide for

extracellular secretion produced easily detectable amounts of this protein (Schillberg *et al.,* 1999). Similarly, retaining human growth hormone in the cytosol of *Nicotiana benthamiana* leaf cells led to protein levels of about 0.01% of TSP, in contrast with concentrations reaching 10% of TSP for the same protein targeted to the apoplast (Gils *et al.,* 2005). Cytosolic accumulation of the human growth factor, for instance, had toxic effects in leaves of *Nicotiana benthamiana* whereas, in contrast, no negative effects were observed for the same protein targeted to the chloroplast (Gils *et al.,* 2005). Likewise, the bovine protease inhibitor (aprotinin) was accumulated at high specific levels when retained in the endoplasmic reticulum (ER) of potato leaf cells (Badri, 2006). Although some recombinant proteins remain stable in the cytosol (e.g. Michaud *et al*., 1998; De Jaeger *et al*., 1999), alternative destinations, such as the chloroplast, the ER or the apoplast, appear to be more appropriate for most proteins.

Targeting to the endoplasmic reticulum

Proteins bearing a signal peptide for cellular secretion first enter the endoplasmic reticulum (ER) via the ER protein translocation channel (Galili *et al*., 1998), and then migrate through this compartment to the golgi apparatus until reaching the extracellular medium (default pathway) or the vacuole, if a vacuolar sorting signal is found in the primary sequence. Recombinant proteins entering the ER may also be retained in this compartment by simple apposition of the tetrapeptide ER retention signal (K /H) DEL (Michaud *et al*., 1998). Plants also accumulate many proteins which do not have a C-terminal retention signal in the ER e.g. prolamins. ER retention signal from γ-zein, a prolamin of maize, has been characterized which is more efficient than KDEL signal (Mainieri *et al*., 2004). Using γ-zein signal, the foreign protein accumulated up to 3.5% of total extractable protein, whereas, it was only

0.5% when KDEL sequence was used. Also, as the protein bodies are insoluble (Mainieri *et al*., 2004), they can be easily purified by centrifugation and it further makes the production of recombinant protein economical. Further, oil bodies originate from ER and store plant seed oils. They are surrounded by a phospholipid monolayer and their membrane is rich in the protein oleosin. A foreign protein could be fused to oleosin with a protease cleavage site in between. This strategy has been used to produce the thrombin inhibitor hirudin from seeds of transgenic *Brassica napus* (Parmenter *et al*., 1995).

Endoplasmic reticulum is considered as a suitable destination for several proteins because of the presence of low abundance of proteolytic enzymes and the presence of molecular chaperones in the ER, together with an oxidizing status favouring disulphide bond formation, helping the protein for proper folding (Nuttall *et al*., 2002; Faye *et al*., 2005). Similar tendencies have been observed for several other proteins of medical or industrial interest. For example human interleukin-4 (Ma *et al*., 2005), the SARS coronavirus S protein antigen (Pogrebnyak *et al*., 2005), the synthetic silk-like protein DR1B (Yang *et al*., 2005) and a recombinant phytase from *Aspergillus niger* (Peng *et al*., 2006). Despite these promising developments, the ER cannot be considered as a suitable destination for all proteins. To be stable or active, a number of clinically useful proteins require late post-translational modifications, such as the formation of complex glycans, the addition of a lipid moiety or the proteolytic removal of a propeptide sequence, which may occur downstream of the ER along the secretory pathway, notably in the Golgi, vacuole or apoplast (Gomord and Faye, 2004; Faye *et al*., 2005). Proteins may exhibit an altered integrity or structural heterogeneity in the ER, as a result of unintended proteolytic processing by ER-resident proteases (Faye *et al*., 2005). For instance, the bovine plasma protein aprotinin expressed in leaves of transgenic potato

showed structural heterogeneity when accumulated in the Endoplasmic reticulum, presumably as a result of the sequential removal of specific amino acids at the N- and C- termini by endogenous peptidases (Badri *et al*., 2005).

Targeting to the peroxisome and nucleus

The recombinant proteins may be sent to these organelles by the inclusion of an appropriate targeting peptide (or localization signal) in the transgene sequence (Daniell *et al*., 2002; Hyunjong *et al*., 2006). For instance a fungal xylanase showed high accumulation levels in *Arabidopsis* leaf tissues sent to peroxisomes using the tripeptide targeting motif SKL (serine-lysine-leucine) grafted at the C-terminus (Hyunjong *et al*., 2006).

Targeting to the chloroplast

Efficient procedures have also been devised to insert transgenes in the chloroplast genome, and then to regenerate transplastomic plant lines accumulating high levels of recombinant protein directly in the chloroplast (Daniell *et al*., 2001b; Maliga, 2002). It eliminates the position effect and results in uniform expression of transgene (Daniell *et al*., 2001a). Human therapeutic protein, interferon gamma was fused to His-tagged GUS and 6% of total soluble protein was expressed in tobacco chloroplasts (Leelavathi and Reddy, 2003). In the tobacco chloroplasts Cry1Ia5 protein accumulated up to 3% of TSP in the leaf tissue (Reddy *et al.,* 2002). De Cosa *et al*., (2001) reported exceptionally high accumulation of Bt Cry2Aa2 proteins (up to 46% of TSP) when the transgene was engineered in chloroplasts. Recently, plastoglobules, the sub-chloroplastic compartments, have, been targeted for recombinant protein accumulation (Vidi *et al*., 2007). Like the oil bodies they are light in weight and contain only a few proteins (Ytterberg *et al*., 2006), which is advantageous for purification. Chlorogen, a biotechnological company,

has adopted the chloroplast transformation as platform technology for the production of foreign proteins in plants. Several transplastomic plant lines have been engineered over the last 10 years for recombinant protein expression, providing very high yields for a number of useful proteins of prokaryotic and eukaryotic origin, including somatotropin, serum albumin, anthrax protective antigen, cholera toxin B subunit and tetanus toxin fragment C (Daniell *et al*., 2001, 2005; Tregoning *et al*., 2003) The chloroplast transformation vector contains two targeting sequences that flank the transgene and insert them through homologous recombination at a particular site.

Chloroplast transformation offers several advantages over nuclear transformation, including uniform transgene expression rates, multiple copies of the transgene in each cell, co-expression of multiple genes from the same construct, minimal gene silencing and minimal transgene escape in the environment owing to the maternal inheritance of chloroplast DNA in several species (Daniell *et al*., 2002). However, a number of endogenous proteases are present in the chloroplast, which can impair the overall stability and accumulation of recombinant proteins (Adam and Clarke, 2002). An interesting example has been provided for the rotavirus VP6 protein, which showed high accumulation rates in chloroplasts of young tobacco leaves, but negligible rates in older leaves (Birch-Machin *et al*., 2004). A similar decline in older tissues was observed for a fungal xylanase (Hyunjong *et al*., 2006) and for the insecticidal Bt toxin Cry2Aa2 (De Cosa *et al*., 2001). Another disadvantage of the chloroplast transgenic system is that plastids do not carry out glycosylation. It is therefore unlikely that chloroplasts could be used to synthesize human glycoproteins in cases in which the glycan-chain structure is crucial for protein activity.

Targeting to the vacuole

Sr. No.	Source organism	Targeting motif	Gene source	Position	Plant species	References
1	Sweet potato	HSRFNPIRLPTTHEPA	Sporamin	N	Sugarcane *Arabidopsis* Tobacco	Gnanasambandam *et al.*, 2004
2	Barley	FADSNPIRPVTDRAAS	Aleurain	N	Tobacco	Holwerdaya *et al.*, 1992
3	Potato	FTSENPIVLPTTCHDDN	22kDa protein	N	Tobacco	Terauchi *et al.*, 2006
4	Potato	FTSQNLIDLPS	Cathepsin D inhibitor	N	Tobacco, *Arabidopsis*	Nakamura *et al.*, 1993
5	Sugarcane vacuolar proteins	IRLPS, ILRLPS, LIRIPL, IRLPS	Legumin, Carboxypeptidase, Aspartic protease II, Trypsin inhibitor	N	Sugarcane	Jackson *et al.*, 2009
6	Barley	NPIR	Aleurain	N	Wheat, Barley	Matsuoka *et al.*, 1991
7	Broad bean	Multiple segments of alpha chains	B-type legumin	N	Heterologues seed	Saalbach *et al.*, 1991
8	Common bean	AFVY	Phaseolin	C	Heterologues leaf	Frigerio *et al.*, 1993

9	Soyabean	Alpha subunit of conglycinin PLSSILRAFY	Alpha subunit of conglycinin	C	Heterologues seed	Nishizawa *et al.*, 2003
10	Soyabean	PFPSILGALY	Beta subunit of conglycinin	C	Heterologues seed	Nishizawa *et al.*, 2004
11	Brazil nut	16 aminoacid long C-terminal stretch	2S albumin	C	Heterologues leaf	Saalbach *et al.*, 1996
12	Tobacco	NGLLVDTM	Chitinase	C	Tobacco	Neuhaus *et al.*, 1991
13	Tobacco	VSGGVWDDSVETNATASLVSEM	Beta 1,3 glucanase	C	Tobacco	Neuhaus *et al.*, 1991
14	Rice	DGMAAILANNQSVSFEGIIESVAELV	Lectin	C	Rice, *Arabidopsis*	Raikhel *et al.*, 1991
15	Barley	VFAEAIAANSTLVAE	Lectin	C	Tobacco	Raikhel *et al.*, 1991
16	Castor	SLLIRPVVPNFN	Ricin	Internal	Heterologues leaf	Frigerio *et al.*, 2007
17	Common bean	Long (32 AA) loop	Phytohemagglutinin	Internal	Heterologues leaf	Schaewen *et al.*, 1993
18	Castor	STGEEVLRMPGDEN	2S albumin	Internal	Heterologues leaf	Brown *et al.*, 2003

Table 4. List of various vacuolar targeting motifs used in plant molecular farming Plant vacuolesnot only maintain the cell turgor but also store proteins and secondary metabolites. There are two distinct types of vacuole in plant cells: lytic (or vegetative) vacuoles, which has an acidic environment rich in hydrolytic enzymes; and protein storage vacuoles, which show a slightly acidic or neutral pH well adapted to protein storage (Robinson *et al*., 2005). Signal sequences that are responsible for targeting the protein to the vacuoles have been identified (Brown *et al*., 2003; Jolliffe *et al*., 2004; Matsuoka and Neuhaus, 1999) but no consensus protein sorting signal has been optimized so far. High accumulation levels have been reported for a numberof recombinant proteins targeted to the vacuole including, synthetic analogue of spider dragline silk protein DP1B in *Arabidopsis thaliana* (Yang *et al.,* 2005), *E. coli* heat-labile enterotoxin B in tobacco (Streatfield *et al.,* 2003), the toxic biotin-binding proteins avidin and streptavidin in tobacco (Murray *et al.,* 2002), *Asgergillus niger* phytase in maize (Arcalis *et al.,* 2004) and a thermo stable glucanase of bacterial origin in barley (Horward *et al.,* 2004).In contrast to protein storing vacuoles, vacuoles of vegetative tissues like leaves have higher hydrolytic activity and recombinant protein targeted would degrade. Therefore, mechanisms or the signals required to store recombinant proteins in the vacuoles need further exploration. A correct *in situ* localization of recombinant proteins bearing a sorting sequence for vacuolar targeting should be undertaken on a systematic basis, considering the species- or tissue-dependent functionality of some sorting signals (Vitale and Hinz, 2005). A very well example of this phenomenon is silk-like protein DP1B expressed in *Arabidopsis* (Yang *et al*., 2005), shows a tissue specific expression of protein reaching 8% of TSP in seed storage vacuoles, and no detectable levels of the same protein could be observed in leaf cell vacuoles. In a similar manner, targeting DP1B to the apoplast provided

good yields in leaves, but poor yields in seeds (Yang *et al*., 2005), again stressing the need for an empirical case-by-case assessment of different tissue and cellular destinations for each protein expressed.A list of various vacuolar targeting motifs used in plant molecular farming is given in table 4. Theimpact of sub cellular targeting on recombinant protein yield in transgenic plant systems was given in table 5.

Yield (fold increase of r-proteins with that of control)									
Protein	**Transformed species**	**Plant organ**	**C**	**ER**	**V**	**A**	**P**	**N**	**Reference**
ScFv anticutinase	*N. tobacum*	Leaf	0	100		1			Schouten *et al*.,1996
ScFv antioxazolone	*N. tobacum*	Leaf		10–20		1			Fielder *et al*., 1997
ScFv anti-oxazolone	*Solanum tuberosum*	Tuber		1		1			Artsaenko *et al*.,1998
ScFv antidihydroflavonol 4-reductase	*Petunia hybrida*	Petal	1	2, 30					DeJaeger *et al*.,1999
BiscFv 2429	*N. tobacum*	BY-2 cells	low	10		1			Fischer *et al*., 1999
FAb MAK33	*Arabidopsis thaliana*	Leaf/seed		1		1			Peeters *et al*., 2001
scFvanti-carcinoembryonic	*N. tobacum*	Leaf		25		1			Stoger *et al*., 2002
Abanti-carcinoembryonic	*N. tobacum*	Leaf		2–6		1			Vaquero *et al*., 2002
Ab 14D9 κ chain	*N. tobacum*	Leaf		8		1			Petruccelli *et*

									al., 2006
Ab 14D9 γ chain				4		1			
Vaccines									
Escherichia coli heat-labile enterotoxin B	*Zea mays*	Seed	1	100	**20,000**	3300	7	21	Streatfield *et al.*,2003
Hepatitis B surface antigen	*N. tobacum*	BY-2 cells	1	1.4		1.8			Sojikul *et al.*, 2003
Japanese cedar pollen allergens	*Oryza sativa*	Seed	0	4–6		1			Takagi *et al.*, 2005
Medical proteins									
Human epidermal growth factor	*N. tobacum*	Leaf	1			10 000			Wirth *et al.*, 2004
Human growth hormone	*N. benthamiana*	Leaf	1			1000	10		Gils *et al.*, 2005

Table 5. Impact of subcellular targeting of recombinant protein yield in transgenic plant systems (modified from Benchabane *et al*., 2008) C-cytoplasm; ER-Endoplasmic reticulum; V-Vacuole; N-Nucleus; A-Apoplast and P- Peroxisomes

Different plant species as platform for molecular farming

The range of plant species amenable to transformation is growing at a phenomenal rate and it is unclear at present which species are optimal for molecular farming. Many factors need to be taken into consideration (Schillberg *et al*., 2003). The factors that are taken into consideration are the total biomass yield, the storage and distribution of the product. Various production platforms have been developed for molecular farming in plants which includes leafy crops (alfalfa, lettuce, *Arabidopsis*, spinach, tobacco), cereals and legumes (barley, maize, pea, pigeon pea, rice, wheat), fruits and vegetables (banana, carrot, potato, tomato, carrots), oil yielding plants (false flax, flax, rape, safflower, soybean, white clover, white mustard) and sugar crops (sugar beet and sugarcane) (Twyman *et al*., 2003 and 2005).

- **Tobacco**

Tobacco have well developed technology for gene transfer and expression, the high biomass yield, the potential for rapid scale-up owing to prolific seed production and the availability of large-scale infrastructure for processing. The demerits of this system include degradation of protein through proteolysis, the presence of toxic alkaloids, and interference of transgene with normal plant metabolism.

- **Cereals and legumes**

Cereals lack the phenolic substances, thereby increasing the efficiency of downstream processing (Ma *et al*., 2003). Legumes, such as alfalfa and soybean, and cereal crops, such as corn and rice, have been considered as ideal candidates for protein production because the

protein can be targeted to accumulate in the seed and the seed can be harvested and stored for an extended amount of time. Alfalfa and soybean produce lower amounts of leaf biomass than tobacco but have the advantage of using atmospheric nitrogen through nitrogen fixation, thereby reducing the need for chemical inputs.

- **Fruits and vegetables**

The main benefit of fruits, vegetables and leafy salad crops is that they can be consumed raw or partially processed, which makes them particularly suitable for the production of recombinant subunit vaccines, food additives and antibodies (Ma *et al*., 2003).

- **Perennial grass**

Perennial grasses like sugarcane provide a 'secure' platform for production of recombinant proteins. Sucrose, the food commodity derived from sugarcane, is sold as a refined crystal that is essentially free of protein, rather than a whole fruit or vegetable. Hence sugarcane producing a pharmaceutical protein was not mixed into the food supply; the food product (refined sucrose) would remain unaffected.

The high yield of recombinant proteins in plant always depends on the properties of protein to be targeted, expressing of these proteins in suitable plant species and targeting these proteins to the right cellular compartment for getting more yields as well as for easy downstream processing. Sugar yielding plants are known for their high biomass production and if proteins can be targeted to the storage tissues there could be a possibility of easy downstream processing as sugarcane juice obtained from sugarcane stem has negligible amount of proteins. This makes sugarcane a better platform for production of commercially important recombinant proteins.

Plant Transformation

Two transformation approaches are commonly used to produce recombinant pharmaceuticals in plants (i) Transient expression and ii) Stable transformation of crop species. There are three major transient expression systems to deliver a gene to plant cells includes delivery of projectiles coated with 'naked DNA' by particle bombardment, infiltration of intact tissue with recombinant *Agrobacterium* (agro infiltration), or infection with modified viral vectors. Stable expressions of transgenes include insertion of genes in the nuclear genome of transgenic plants by two general methods *Agrobacterium*-mediated transformation and particle bombardment.

Transient expression in plants

The transient production platform is perhaps the fastest and the most convenient production platform for plant molecular farming (Rybicki, 2010). The systems, which are mainly used for quick validation of expression constructs, are routinely used for the production of considerable amounts of proteins within a few weeks (Vezina *et al.,* 2009). The following methods are commonly used for transient expression in plants.

Agro infiltration

The agro infiltration method, which was developed by Kapila *et al.,* (1997), involves infiltration of a suspension of recombinant *Agrobacterium tumefaciens* into tobacco leaf tissue, which facilitates the transfer of T-DNA to a very high percentage of the cells, where it expresses the transgene at very high levels without stable transformation, as in the case of transgenic crops. This method has now been developed into a very rapid, high-yielding transient expression strategy for producing clinical grade bio-pharmaceuticals (Vézina *et al.,* 2009; Pogue *et al.,* 2010; Regnard *et al.,* 2010).

Viral infection

The viral infection method is dependent on the ability of plant viruses such as tobacco mosaic virus (TMV) and potato virus X (PVX) to be used as vectors to deliver foreign genes into plants, without integration (Porta and Lomonossoff, 2002). Both expression platforms infect tobacco plants and then transiently express a target protein in the plant tissue. Using this expression method Varsani *et al.,* (2006) were able to successfully obtain protein yield as high as 17% of the total protein. The main drawback of this system was a tendency to lose the foreign insert during spread of the virus throughout the plant and the potential environmental issues associated with the presence of infectious recombinant viral particles.However, in the transient expression system the recombinant protein has to be processed immediately to prevent tissue degradation and protein instability.

Stable nuclear transformation

Stable nuclear transformation involves the incorporation of a foreign gene of interest into the nuclear genome of the plant, thereby altering its genetic makeup, and leading to the expression of the transgene. The stable nuclear transformation has produced most of the recombinant proteins till date. The method has been used to accumulate protein in the dry seeds of cereals, which allows long term storage of the seed at room temperature without degradation of the protein (Horn *et al.,* 2004). These include the following methods

Transformation via particle bombardment

Particle bombardment is one of the most widely used plant transformation method and has been applied to a broad range of species, especially monocots. The process involves the introduction of genetic material into intact cells and tissues through the use of high velocity micro projectiles. The high velocity microprojectiles are used to carry DNA being ‘shot’ into cells which represents a type of biological

ballistics, hence the term 'biolistics' (Sanford, 1990). Efficient transformation of sugarcane through particle bombardment has been achieved by using embryogenic callus cultures (Bower and Birch, 1992). Based on acceleration of microscopic tungsten or gold particles coated with DNA can be propelled into practically all kind of tissues like callus (Ritala *et al.,* 1994) suspension culture (Gordon-Kamm *et al.,* 1990) inflorescences (Barcelo *et al.,* 1994) shoot apices (Zhong *et al.,* 1996) microspores (Jahne *et al.,* 1994) leaves (Tomes *et al.,* 1990) roots (Seki *et al.,* 1991) or pollen grains (Stoeger *et al.,* 1995) and stable transgenic cereal plants were reported for wheat (Vasil *et al.,* 1992), oat (Somers *et al.,* 1992), barley (Wan and Lemaux, 1994) as well as for rye (Castillo *et al.,* 1994). Although particle bombardment can be used to test recombinant protein stability before stable transformation, it is unsuitable for the expression of larger amounts of foreign proteins because of low efficiency, reproducibility and regeneration potential was also limited for long term callus cultures (Christou, 1996).

***Agrobacterium* mediated transformation**

Agrobacterium tumefaciens is a gram-negative soil bacterium used to transfer foreign genes to plants. The merits include integration of the small copy number of T-DNA into plant chromosomes, and stable expression of transferred genes.*Agrobacterium*-mediated gene transfer offers potential advantages, including preferential integration of the transgene into transcriptionally active regions of the chromosome (Koncz *et al*., 1989) with exclusion of vector DNA (Hiei et al., 1997), unlinked integration of co-transformed T-DNAs (McKnight *et al.,* 1987), transfer of large DNA fragments (Liu *et al*., 1999). This method is normally used for the transformation of dicot species (De Block *et al*., 1985) and has been used in tobacco, alfalfa, pea, tomato and potato (Ma *et al*., 2003). Monocots are more difficult to transform using *Agrobacterium* mediated

transformation. However, monocots have also been transformed by *Agrobacterium*, and the technology has been optimized for selected model varieties. *Agrobacterium* has been used to transform the plants of Gladiolus genus (Graves and Goldman, 1987), maize (Ishida *et al*., 1996), barley (Wu *et al*., 1998), rice (Dong *et al*., 1996; Rashid *et al*., 1996), Asparagus (Hernalsteens *et al*., 1984), banana (Sreeramanan *et al*., 2009; May *et al*., 1995). *Agrobacterium*-mediated DNA transfer to sugarcane meristems has been attempted and optimized to give better efficiency (Arencibia *et al*., 1995, EnroAquez-Obregoan *et al*. 1997). Consequently, herbicide-resistant sugarcane plants have been produced through *Agrobacterium* transformation (Enriquez Obregon *et al*., 1998).

The interaction of two compatible plasmids, one containing the *vir*-region, the other carrying the T-DNA on a wide host-range replicon forms a binary vector system (Hoekema *et al*., 1983). A binary vector system is generally used for *Agrobacterium* mediated transformation. Agrobacterium contain the T-DNA, which is stably integrated into the plant genome. Apart from the T-DNA another region of the Ti-plasmid-called the vir-region, is essential for tumour induction. Transfer of the plasmid into an *A. tumefaciens* strain harbouring the plasmid with the vir-region allows introduction of the manipulated T-DNA into plant cells. Production of recombinant proteins in transgenic plants was initially based on integration of a target gene into the nuclear genome, expression of *virurence* (*vir*) genes from the Ti plasmid of *Agrobacterium tumefaciens* are enhanced by several experimental factors, including phenolic compounds (Spencer and Towers, 1988), acetosyringone concentration from 100 to 200 µM, sugars (Shimoda *et al*., 1990) and pH of co-cultivation media, ranging from pH 5.4 to 5.6 (Mondal *et al*., 2001). The T-DNAs insertion is random into the genome (Ambros, 1986; Wallroth, 1986) and remains stable in the original insertion site through multiple generations (Krysan *et al*., 1999). Biolistic

transformation invariably leads to the genomic insertion of multiple copies of the transgene cassette, which in turn can lead to gene silencing. Hence, *Agrobacterium*-mediated transformation may be another alternative to reduce somaclonal variation and for overcoming gene silencing.

Optimization of transgene expression in host plant

Plant as expression systems offer many benefits over conventional systems. However, there are still many challenges that remain to be overcome to use plants as main stream production platform. For the development of plant-based production platform, one needs to optimize the expression level of a recombinant protein.

Optimization of promoters and terminators

In order to optimize transcript expression, the general strategy is to use strong and constitutive promoters, such as the cauliflower mosaic virus 35S RNA promoter (CaMV 35S) and maize ubiquitin-1 promoter (ubi-1), for dicots and monocot, respectively (Fischer *et al.,* 2004). Organ and tissue-specific promoters are also being used to drive expression of the transgenes in the tissue or organ like the tuber, the seed and the fruit (He *et al.,* 2008). This tissue specific expression therefore prevents accumulation of the recombinant protein in the vegetative organs, which might negatively affect plant growth and development. Additionally, inducible promoters, whose activities are regulated by either chemical or external stimulus, may equally be used to prevent the lethality problem (Corrado and Karali, 2009), as it is being used in cell suspension cultures (Nara *et al.,* 2000; Peebles *et al.,* 2007). Besides, transcription factors can be used as boosters for the promoters to further enhance the expression level of the transgenes (Yang *et al.,* 2001). Moreover, it has been recently found that the terminator of the heat-shock gene of *Arabidopsis*

thaliana shows an increase in transcription of a foreign gene by four times (Nagaya *et al.,* 2010).

Furthermore, the expression constructs can also be designed to ensure transcript stability and translational efficiency. This involves, for instance, the removal of the native 5′ and 3′ untranslated regions from the foreign gene and introducing 5′ untranslated leader sequence of the tobacco mosaic virus RNA, rice polyubiquitin gene RUB13, alfalfa mosaic virus or tobacco etch virus in the expression construct, all of which have been separately shown to significantly enhance the expression levels of transgenes (Lu *et al.,* 2008; Sharma *et al.,* 2008). 5′ UTR of rice poly ubiquitin gene RUBI3 along with its promoter was reported to enhance the expression of GUS at mRNA level as well as translational level suggesting that 5′ UTR plays an important role in gene expression (Lu *et al.,* 2008; Samadder *et al.,* 2008). The untranslated leader sequences of alfalfa mosaic virus mRNA 4 or tobacco etch virus have been found to enhance the transgene expression by several folds due to enhanced translational efficiency of transcripts (Datla *et al.,* 1993; Gallie *et al.,* 1995). In addition to these leader sequences, the expression cassette design such that the false AU-rich sequences in the 3′ untranslated regions that may act as splice sites are removed or modified, to ensure transcript stability (Mishra *et al.,* 2006). Besides, the transcript stability can be ensured by co-expressing the gene of interest and a suppressor of RNA silencing (Voinnet *et al.,* 2003). It is also established that each organism exhibits biased codon usage, such that it might be important to adapt the coding sequence of the heterologous gene to that of the host plant in order to optimize translation efficiency (Lienard *et al.,* 2007). In this regard, the translational start-site of the heterologous protein is modified to match with the Kozak consensus for plants (Kawaguchi and Bailey-Serres, 2002) or by using the sequence GCT TCC TCC after initiation codon, or ACC or ACA before it (Sharma and Sharma, 2009).

Optimization of codon usage

The codon modification, however, should be empirically determined rather than predicted because of the variation in the levels of transgene expression in the same system, using the same construct (Rybicki, 2009). To this end, codon combinations (A/G)(a/c)(a/g) AUG and (A/G)(u/C)(g/C) AUG have been reported to be optimum for enhanced translational activity in Arabidopsis and rice, respectively (Sugio *et al.,* 2010). This variation in the expression of the transgene may be due to position effect, copy number of the transgene or gene silencing. With respect to position effect, expression cassettes can now be designed to the nuclear matrix attachment regions (MAR), which are regulatory sequences that ensure the placement of the transgene in suitable regions for mobilizing transcription factors to the promoters (Streatfield, 2007). Besides, the problem of position effect can be avoided by targeting the transgene into the plastids (Cardi *et al.,* 2010). For optimizing the generation of single-copy transgenics, the strategies that have been used include the use of specific genetic elements, including the cAMP response elements (CREs), for co transfer with transgene in the T-DNA (De Paepe *et al.,* 2009). Additionally, a new technology, which consists of the construction of genetically autonomous artificial mini chromosomes, has been described as providing infinite possibilities with several enormous advantages including gene stability; owing to the absence of gene silencing and position effect (Ananiev *et al.,* 2009).

Downstream processing of recombinant proteins

Along with high level of transgene expression to provide good yields in plant-based production system, efficient recovery of the recombinant proteins must also be optimized. The goal and the general steps for downstream processing are similar between plant and other

expression systems. The goal is to recover the maximal amount of highly purified target protein with the minimal number of steps and at the lowest cost. The basic steps for downstream processes include tissue harvesting, protein extraction, purification, and formulation. Plants contain several undesired molecules (proteins, oils, phenolic compounds etc.) that must be removed during purification of the recombinant protein. The processing of plant tissue for the recovery of recombinant proteins generally includes fractionation of plant tissue, extraction of recombinant protein in aqueous medium and purification. Several strategies have been developed to improve the downstream processing of plant produced recombinant proteins.

Use of affinity tags for purification

The use of affinity tags with the desired protein is a powerful approach for recombinant protein recovery, but the tag needs to be removed from the final product. Several affinity tags have been described for recombinant protein purification (Terpe, 2003). N-terminal hexa histidine tagged ricin B (HIS-RTB), the lactin subunit of ricin, was expressed in tobacco. The recombinant HIS-RTB was purified from tobacco leaves using lactose affinity column, and was found to be functional (Reed *et al.,* 2005). His-tagged amarantin was expressed in tobacco and the presence of His tag was used for single-step purification of recombinant protein using immobilized metal-ion affinity chromatography (Valdez-Ortiz *et al.,* 2005). The effects of three affinity tags, i.e. eight amino acid tag StrepII, His6 and 181-amino acid Tandem Affinity Purification (TAP) tag were studied for the purification of recombinant membrane anchored protein kinase. The protein purified using His6 tag was of low purity whereas the recombinant proteins having TAP or StrepII tag were purified to homogeneity. While StrepII-tag purification achieved high yield and purity which was comparable to that obtained with TAP-tag, it was considerably easier and faster (Witte *et*

*al.,*2004). In addition to aid in the purification of recombinant protein from plants (Conley *et al.,* 2009; Joensuu *et al.,* 2009), elastin like polypeptide tag has also been found to enhance expression of the protein target (Conley *et al*., 2009). The possibilities of using histamine, tryptamine, phenylamine or tryamine as affinity tags for affinity purification of recombinant proteins have also been evaluated (Platis and Labrou, 2008). The protein splicing elements (inteins) that can catalyse self-cleavage (Liu, 2000) have also been utilized in protein purification purposes. The SMAP 29, a mammalian antimicrobial peptide (Bagella *et al.,* 1995), was expressed with intein fusion tag in tobacco plants.

Use of non-chromatographic methods

Mostly the non-chromatographic method involves capturing and partial purifying of oleosin-fusion proteins by centrifugation (Van Rooijen and Moloney, 1995). These fusions have oleosin proteins as N-terminal fusion partners that allow *in vivo* targeting and/or post-extraction capture of the target protein on the surface of oil bodies (Boothe et al., 2010). Purification of oil bodies with attached fusions is done by several washing steps in aqueous solutions using centrifugation. Attached recombinant protein can be released from oil bodies by proteolytic cleavage of the oleosin-target protein linkage or by elution in the case of affinity bound proteins. The affinity binding approach is exemplified by constructs developed recently that contain an N-terminal anti-oleosin single chain antibody (scFv) as the fusion partner. Following their elution from oil bodies these fusions can be cleaved either chemically (e.g. acidic cleavage) or enzymatically to release the recombinant protein (Boothe et al., 2010 and Nykiforuk et al., 2011). The potential downside of these two strategies is the need to cleave the recombinant protein from the oleosin or scFv fusion partner. The cleavage precision and efficiency,

whether chemical or enzymatic, is typically less than 70% and results in reduced product yield. Subsequent purification of released recombinant proteins from respective fusions is done by standard adsorption chromatography. Other equally good examples of non-chromatographic purification methods utilized unique protein properties (size, hydrophobicity or stability) and host system properties to accomplish enhanced separation efficiency. Lee and Forciniti (2010) explored the use of aqueous two-phase (PEG/salt) partitioning as a sole recovery and purification method of non-glycosylated mAb expressed in corn seed. By manipulating the system composition, pH, and ionic strength they managed to partially purify the antibody in a three-stage process. The first two stages were typical two-phase partitioning with the target protein concentration enriched in the bottom (aqueous salt phase) and host impurities in the top (PEG) phase. These two-extraction protocols resulted in a rather modest mAb purification of 1.3- and 1.4-fold, respectively. The third stage consisted of mAb precipitation at the two-phase interface and resulted in almost 10-fold purification. Overall, the three-stage processes delivered 72% pure mAb with 49% yield. Clearly, an additional adsorption step, most likely affinity chromatography, would be needed to purify the antibody for biopharmaceutical applications. Aspelund and Glatz (2010) demonstrated purification of recombinant collagen from low pH corn extracts by cross-flow filtration. Diafiltration of corn endosperm extracts at pH 3.1 by using 100-kDa MWCO membranes removed 96% of host protein and resulted in 89% pure collagen. Improved purification of collagen was achieved by protein precipitation of endosperm extracts with sodium chloride at pH 2.1. Thus, the unique composition of endosperm extract and molecular properties of collagen (high molecular weight and stability at low pH) allowed the development of this extremely attractive and inexpensive purification scheme.

Although much progress has been made over the past seven years, downstream processing still requires further attention and technological breakthroughs to fulfil the long anticipated goal of low manufacturing cost for plant-derived recombinant proteins. However, improvements in downstream processing alone will not suffice; high protein expression levels and assured product fidelity are also necessary.

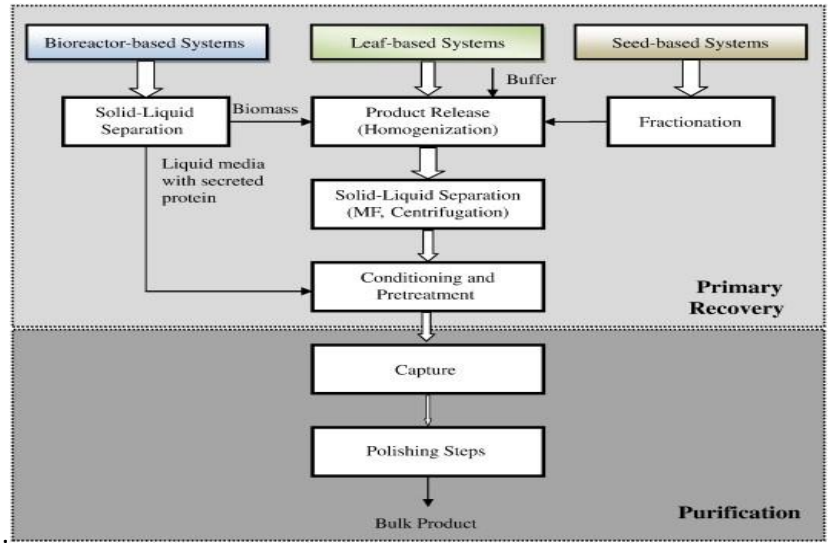


Fig 2. Downstream processing of recombinant proteins from ioreactor-based, leaf-based, and seed-based systems

Biosafety and regulatory issues

Concerns have been raised about the safety of GM foods in relation to environment and human health. Although there is no scientific evidence that current modified foods involve any new or magnified risks, certain environmentalists and consumers are still not convinced (Smyth and Phillips, 2003). The general public concern about the potential health and environmental risks associated with the PMF crops (not the products) is being viewed at two levels; in that, not only

are they engineered to accumulate all sorts of proteins with medicinal properties at very high concentration, which may affect the host plants (Badri *et al.,* 2009), but also that their biologically active products are meant to elicit physiological responses in humans and in animals (Spök, 2007; Spök et al., 2008). Additionally, there are specific concerns, including the lack of communication among the regulatory bodies involved in research, biosafety and trade, further hampers the developments in this field (Ramessar *et al.,* 2008b). The regulation of pharmaceutical crops is in its infancy and there are several challenges ahead for the regulatory agencies. There is a lot of pressure from pharmaceutical industry, food industry, environmental and consumer organizations against GM crops and regulations are strict and turn out to be very costly. There is a requirement to regulate the pharmaceutical crops on case by case basis. The regulatory challenges posed ahead for the molecular farming and how they are different from those for first generation transgenic crops, have been reviewed recently (Spok, 2007; Spok *et al.,* 2008). The strategies used for risk assessment need to be reviewed. The most important issue is to segregate the GM crops from non-GM crops to prevent intermixing. A variety of approaches including physical containment as well as genetic strategies like seed sterility, maternal inheritance, male sterility, selective elimination by engineering sensitivity to chemicals, etc. have been postulated to address this question (Howard and Donnelly, 2004; Lee and Natesan, 2006; Lin *et al.*, 2008) and threshold limits of accidental contaminations have been suggested. Like for the GM food and feed crops, several regulations are being developed, to increase the biosafety of the plant bioreactors, even though, one knows that, there is no fool-proof system, as there might be some elements of human errors and natural accidents, which are beyond control.

Concluding remarks

Key to the success of transgenic plants derived products in the future will be based on the level of expression achieved. It is very important with regard to economics, because it affects cost of growing, processing, extraction purification and waste disposal. It is clear that attempts would be made towards higher levels of expression. If the technical hurdles can be overcome, soon it might be possible to make protein-based pharmaceuticals available to needy at affordable cost. Most of all, the overall prospects of plant-based molecular farming industry will depend on improved public perception of the technology and the products, especially as the industry improves its level of compliance with the regulations, and as there are many successful clinical trials and approvals of the first set of these plant derived human pharmaceuticals The full realization and impact of the aforementioned developments, however, depends not only on consistent, successful and innovative research and developmental activities, but also on a favorable regulatory climate and public acceptance. Overall production of plant-derived biologics is going to be an important methodology for the future.

References

1. Adam Z, Clarke AK. 2002. Cutting edge of chloroplast proteolysis. *Trends Plant Sci* 7(10): 451-456.

2. Akama K, Kanetou J, Shimosaki S, Kawakami K, Tsuchikura S, Takaiwa F. 2009. Seed-specific expression of truncated OsGAD2 produces GABA-enriched rice grains that influence a decrease in blood pressure in spontaneously hypertensive rats. *Transgenic Res* 18(6): 865-876.

3. Ambros P, Matzke AJM, Matzke MA .1986. Localization of Agrobacterium rhizogenes T-DNA in plant chromosomes by in siteu hybridisation. EMBO 5, 2073- 2077.

4. Ananiev EV,Wu C,ChamberlinMA,Svitashev S,Schwartz C,Gordon-KammW, et al. 2009. Artificial chromosome formation in maize (Zea mays L.). *Chromosoma;*118:157–77.

5. Anderson AJ, Dawes EA. 1990. Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol Rev* 54(4): 450-472.

6. Altmann F, Kolarich D, Drakakaki G, Fischer R, Christou P, Stoger E. 2004. Unexpected deposition patterns of recombinant proteins in post-endoplasmic reticulum compartments of wheat endosperm. *Plant Physiol* 136(3): 3457-3466

7. Arencibia AD, Carmona E R, Tellez P, Chan MT, Yu, S M, Trcijillo, L E and Oramas P. 1998. An efficient protocol for sugarcane (Saccharum spp. L.) transformation mediated by *Agrobacterium tumefaciens*. *Transgenic Res.* 7: 213-222.

8. Artnzen CJ. 1997. Edible vaccines. *Public Health Rep* 112(3): 190-197.

9. Aspelund MT, Glatz CE. 2010. Purification of recombinant plant-made proteins from cornextracts by ultrafiltration. *J Membr Sci* 353:103–10.

10. Badri, AM, Goulet, C, Rivard, D, Li, H, Coenen, K. and Michaud, D. 2006 A SELDI-TOF/MS procedure for the rapid detection and quantitation of recombinant proteins in transgenic plants. In: *Proceedings of the Conference on Plant-Made Pharmaceuticals 2005* (Arcand, F., ed.), pp.82–83. Québec, QC: Society for Moleculture.
11. Bagella L, Scocchi M, Zanetti M. 1995. cDNA sequences of three sheep myeloid cathelicidins. FEBS Lett; 376:225–8.
12. Barcelo P, Hagel C, Becker D, Martin,A. and Lörz, H. 1994. Transgenic cereal (tritordeum) plants obtained at high efficiency by microprojectile bombardment of inflorescence tissue. *Plant J.* 5:583- 592.
13. Barta A, Sommergruber K, Thompson D, Hartmuth K, Matzke MA, Matzke AJ. 1986. The expression of a nopaline synthase - human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. *Plant Mol Biol* 6(5): 347-357.
14. Birch-Machin I, Newell CA, Hibberd JM, Gray JC. 2004. Accumulation of rotavirus VP6 protein in chloroplasts of transplastomic tobacco is limited by protein stability. *Plant Biotechnol J* 2(3): 261-270.
15. Boehm R. 2007. Bioproduction of therapeutic proteins in the 21st century and the role of plants and plant cells as production platforms. *Ann N Y Acad Sci* 1102: 121-134.
16. Bohmert K, Balbo I, Kopka J, Mittendorf V, Nawrath C, Poirier Y, Tischendorf G, Trethewey RN, Willmitzer L. 2000. Transgenic Arabidopsis plants can accumulate polyhydroxybutyrate to up to 4% of their fresh weight. *Planta* 211(6): 841-845.
17. Bower, R and R.G. Birch. 1992. Transgenic sugarcane plants via microprojectile bombardment. *Plant J.* 2:409-416.

18. Brown JC, Jolliffe NA, Frigerio L, Roberts LM. 2003. Sequence-specific, Golgi-dependent vacuolar targeting of castor bean 2S albumin. *Plant J* 36(5): 711-719.
19. Cahoon EB, Shockey JM, Dietrich CR, Gidda SK, Mullen RT, Dyer JM. 2007. Engineering oilseeds for sustainable production of industrial and nutritional feedstocks: solving bottlenecks in fatty acid flux. *Curr Opin Plant Biol* 10(3): 236-244.
20. Cardi T, Lenzi P, Maliga P. 2010. Chloroplasts as expression platform for plant-produced vaccines. *Expert Rev Vaccin* 9:893–911.
21. Castillo AM, Vasil V and Vasil IK. 1994. Rapid production of fertile transgenic plants of rye (Secale cereale L.) *BioTechnology* 12: 1366–1371
22. Conley AJ, Joensuu JJ, Jevnikar AM, Menassa R, Brandle JE. 2009. Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. *Biotechnol Bioeng* 103:562–73
23. Conrad U, Fiedler U. 1998. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Mol Biol* 38(1-2): 101-109.
24. Corrado G, Karali M. 2009. Inducible gene expression systems and plant biotechnology. *Biotechnol Adv* 27:733–43.
25. Dacks JB, Peden AA, Field MC. 2009. Evolution of specificity in the eukaryotic endomembrane system. *Int J Biochem Cell Biol* 41(2): 330-340.
26. Daniell H .2002. Molecular strategies for gene containment in transgenic crops. *Nat Biotechnol* 20: 581–586

27. Daniell H, Kumar S, Dufourmantel N. 2005. b Breakthrough in chloroplast genetic engineering of agronomically important crops. *Trends Biotechnol* 23: 238–245
28. Daniell H, Lee SB, Panchal T, Wiebe PO (2001. a) Expression of the native cholera toxin B subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. J Mol Biol 311 1001–1009
29. Daniell H, Muthukumar B, Lee SB (2001. b) Marker free transgenic plants: engineering the chloroplast genome without the use of antibiotic selection. Curr Genet 39 109–116
30. Daniell H, Streatfield SJ, Wycoff K. 2001. Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci* 6(5): 219-226.
31. Daniell H. 2006. Production of biopharmaceuticals and vaccines in plants via the chloroplast genome. *Biotechnol J* 1(10): 1071-1079.
32. Daniell H. 2006. Production of biopharmaceuticals and vaccines in plants via the chloroplast genome. *Biotechnol J* 1(10): 1071-1079.
33. Datla RS, Bekkaoui F, Hammerlindl JK, Pilate G, Dunstan DI, Crosby WL. 1993. Improved highlevel constitutive foreign gene expression in plants using an AMV RNA4 untranslated leader sequence. *Plant Sci* 94:139–49.
34. De Cosa B, Moar W, Lee SB, Miller M, Daniell H. 2001. Overexpression of the Bt cry2Aa2 operon in chloroplasts leads to formation of insecticidal crystals. *Nat Biotechnol* 19(1): 71-74.
35. De Jaeger G, Buys E, Eeckhout D, De Wilde C, Jacobs A, Kapila J, Angenon G, Van Montagu M, Gerats T, Depicker A. 1999. High level accumulation of single-chain variable fragments in the cytosol of transgenic Petunia hybrida. *Eur J Biochem* 259(1-2): 426-434.
36. De Jaeger G, Scheffer S, Jacobs A, Zambre M, Zobell O, Goossens A, Depicker A, Angenon G. 2002. Boosting heterologous protein

production in transgenic dicotyledonous seeds using Phaseolus vulgaris regulatory sequences. *Nat Biotechnol* 20(12): 1265-1268.

37. De Paepe A, De Buck S, Hoorelbeke K, Nolf J, Peck I, Depicker. 2009. A high frequency of single copy T-DNA transformants by floral dip in CRE-expressing Arabidopsis plants. *Plant J* 59:517–27.
38. Dong J, Teng W, Buchholz W G and Hall TC. 1996. Agrobacterium-mediated transformation of Javanica rice. *Molecular Breeding* 2:267-276.
39. Enriquez-Obrego´n GA, Vaquez-Padro´n RI, Prieto-Samso´nov D L, de la Riva GA and Selman-Housein G. 1998. Herbicide-resistant sugarcane (Saccharum officinarum L.) plants by Agrobacterium-mediated transformation. *Planta* 206:20-27.
40. Faye L, Boulaflous A, Benchabane M, Gomord V, Michaud D. 2005. Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* 23(15): 1770-1778.
41. Faye L, Boulaflous A, Benchabane M, Gomord V, Michaud D. 2005. Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* 23(15): 1770-1778.
42. Faye L, Gomord V. 2010. Success stories in molecular farming-a brief overview. *Plant Biotechnol J* 8(5): 525-528.
43. Fischer R, Emans NJ, Twyman RM, Schillberg S. 2004. Molecular farming in plants: technology platforms. In: Goodman RB, editor. Encyclopedia of plant and crop science. *New York: Marcel Dekker*. p. 753–6
44. Fischer R, Twyman RM, Schillberg S. 2003. Production of antibodies in plants and their use for global health. *Vaccine* 21(7-8): 820-825.

45. Galili G, Galili S, Lewinsohn E, Tadmor Y. 2002. Genetic, molecular and genomic approaches to improve the value of plant foods and feeds. *Crit Rev Plant Sci* 21:167–204
46. Gallie DR, Tanguay RL, Leathers V. 1995. The tobacco etch viral 5′ leader and poly (A) tail are
47. functionally synergistic regulators of translation. *Gene* 165:233–8.
48. Gils M, Kandzia R, Marillonnet S, Klimyuk V, Gleba Y. 2005. High-yield production of authentic human growth hormone using a plant virus-based expression system. *Plant Biotechnol J* 3(6): 613-620.
49. Goldstein DA, Thomas JA. 2004. Biopharmaceuticals derived from genetically modified plants. *QJM* 97(11): 705-716.
50. Gomord V, Denmat LA, Fitchette-Laine AC, Satiat-Jeunemaitre B, Hawes C, Faye L. 1997. The C-terminal HDEL sequence is sufficient for retention of secretory proteins in the endoplasmic reticulum (ER) but promotes vacuolar targeting of proteins that escape the ER. *Plant J* 11(2): 313-325.
51. Gordon-Kamm, WJ, Spencer TM, Mangano ML, Adams TR, Daines RJ, Start WG, O'Brien JV, Chambers SA, Adams Jr WR, Willetts NG, Rice TB, Mackey, CJ, Krueger, RW, Kausch, AP and Lemaux, PG. 1990. Transformation of maize cells and regeneration of fertile transgenic plants. *Plant Cell* 2:603-618.
52. Goulet, C. and Michaud, D. 2006. Degradation and stabilization of recombinant proteins in plants In: *Floriculture, Ornamental and Plant Biotechnology*, Vol. IV (Taxeira da Silva, J., ed.), pp. 35–40. London: Global Science Books.
53. Graves AC. and Goldman SL 1986. The transformation of Zea mays seedlings with Agrobacterium tumefaciens: Detection of T-DNA specific enzyme activity. *Plant MOI. Biol.* 7, 43-50.

54. Hahn JJ, Eschenlauer AC, Sleytr UB, Somers DA, Srienc F. 1999. Peroxisomes as sites for synthesis of polyhydroxyalkanoates in transgenic plants. *Biotechnol Prog* 15(6): 1053-1057.
55. He Z, Jiang XL, Qi Y, Di QL. 2008.Assessment of the utility of the tomato fruit-specific E8 promoter for driving vaccine antigen expression. *Genetica* 133:207–14
56. Hernalsteens JP, Thia-Toong L, Schell J, Van Montagu M. An Agrobacterium-transformed cell culture from the monocot Asparagus officinalis. *EMBO J.* 1984 Dec 20;3(13):3039–3041
57. Hiatt A, Cafferkey R, Bowdish K. 1989. Production of antibodies in transgenic plants. *Nature* 342(6245): 76-78.
58. Hiei, Y, Komari, T and Kubo, T. 1997. Transformation of rice mediated by Agrobacteriumtumefaciens. *Plant Molecular Biology* 35: 205-218
59. HoekemaA PR, Hirsh P JJ, Hooykaas and Schilperoort R A. 1983. A binary plant vector strategy based on separation of *vir*-and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature* 303:179-180.
60. Hood EE, Witcher DR, Maddock S, Meyer T, Baszczynski C, Bailey M, et al. 1997. Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Mol Breed* 3:291–306.
61. Horn ME, Woodard SL, Howard JA. 2004. Plant molecular farming: systems and products. *Plant Cell Rep* 22(10): 711-720.
62. Howard JA, Donnelly KC. 2004. A quantitative safety assessment model for transgenic protein products produced in agricultural crops. *J Agric Environ Ethics* 17:545–58.
63. Howard JA, Donnelly KC. 2004. A quantitative safety assessment model for transgenic protein products produced in agricultural crops. *J Agric Environ Ethics* 17:545–58.

64. Hyunjong B, Lee DS, Hwang I. 2006. Dual targeting of xylanase to chloroplasts and peroxisomes as a means to increase protein accumulation in plant cells. *J Exp Bot* 57(1): 161-169.
65. Ishida Y, Saito H, Ohta S, Hiei Y, Komari T and Kumashiro T. 1996 .High efficiency transformation of maize (Zea mays L.) mediated by Agrobacterium tumefaciens. *Nat. Biotec.* 14(6):745-750.
66. Jahne A, Becker D, Brettschneider R and Lörz H. 1994. Generation of transgenic, microspore-derived fertile barley. *Theor. Appl. Genet*. 89: 525-533.
67. Joensuu JJ, Brown KD, Conley AJ, Clavijo A, Menassa R, Brandle JE. 2009. Expression and purification of an anti-foot-and-mouth disease virus single chain variable antibody fragment in tobacco plants. *Transgenic Res* doi:10.1007/s11248-009-9257-0.
68. John ME, Keller G. 1996. Metabolic pathway engineering in cotton: biosynthesis of polyhydroxybutyrate in fiber cells. *Proc Natl Acad Sci U S A* 93(23): 12768-12773.
69. Jolliffe NA, Brown JC, Neumann U, Vicre M, Bachi A, Hawes C, Ceriotti A, Roberts LM, Frigerio L. 2004. Transport of ricin and 2S albumin precursors to the storage vacuoles of Ricinus communis endosperm involves the Golgi and VSR-like receptors. *Plant J* 39(6): 821-833.
70. Kapila J, De Rycke R, van Montagu M, Angenon G. 1997.An agrobacterium-mediated transient gene expression system for intact leaves. *Plant Sci* 122:101–8.
71. Kawaguchi R, Bailey-Serres J. 2002. Regulation of translational initiation in plants. *Curr Opin Plant Biol* 5:460–5.
72. Khandelwal A, Sita GL, Shaila MS. 2003. Expression of hemagglutinin protein of rinderpest virus in transgenic tobacco and immunogenicity of plant-derived protein in a mouse model. *Virology* 308(2): 207-215.

73. Kinney AJ. 2006. Metabolic engineering in plants for human health and nutrition. *Curr Opin Biotechnol* 17(2): 130-138.
74. Koncz C, Martin N, Mayerhofer R, et al. 1989. High-frequency T-DNA-mediated gene tagging in plants. Proceedings of the National Academy of Sciences, 86, 8467–8471.
75. Krysan PJ, Young JC, and Sussman MR 1999. T-DNA as an insertional mutagen in Arabidopsis. *Plant Cell* 11, 2283-2290.
76. Lee D, Natesan E. 2006. Evaluating genetic containment strategies for transgenic plants. *Trends Biotechnol* 24:109–14.
77. Lee JW, Forciniti D. 2010. Purification of human antibodies from transgenic corn using aqueous two-phase systems. *Biotechnol Progr* 26:159–67
78. Leelavathi S, Reddy VS. 2003. Chloroplast expression of His-tagged GUS-fusions: a general strategy to overproduce and purify foreign proteins using transplastomic plants as bioreactors. *Mol Breed* 11:49–58.
79. Liénard D, Sourrouille C, Gomord V, Faye L. 2007. Pharming and transgenic plants. *Biotechnol Ann Rev* 13:115–47.
80. Liu XQ. 2000. Protein-splicing intein: Genetic mobility, origin, and evolution. *Annu Rev Genet*. 34:61-76.
81. Lossl A, Bohmert K, Harloff H, Eibl C, Muhlbauer S, Koop HU. 2005. Inducible trans-activation of plastid transgenes: expression of the R. eutropha phb operon in transplastomic tobacco. *Plant Cell Physiol* 46(9): 1462-1471.
82. Lu J, Sivamani E, Azhakanandam K, Samadder P, Li X, Qu R. 2008. Gene expression enhancement mediated by the 5′ UTR intron of the rice rubi3 gene varied remarkably among tissues in transgenic rice plants. *Mol Genet Genomics* 279:563–72.
83. Ma JK, Drake PM, Christou P. The production of recombinant pharmaceutical proteins in plants. Nat Rev Genet 2003;4:794–805.

84. Ma S, Wang A .2012. Molecular Farming in Plants: An Overview, in Molecular Farming in Plants: *Recent Advances and Future Prospects*, Springer 1-20.
85. Mainieri D, Rossi M, Archinti M, Bellucci M, De Marchis F, Vavassori S, Pompa A, Arcioni S, Vitale A. 2004. Zeolin. A new recombinant storage protein constructed using maize *γ*-zein and bean phaseolin. *Plant Physiol* 136**:** 3447–3456
86. Maliga P. 2002. Engineering the plastid genome of higher plants. *Curr Opin Plant Biol* 5(2): 164-172.
87. Matsumoto K, Wassarman KM, Wolffe AP. 1998. Nuclear history of a pre-mRNA determines the translational activity of cytoplasmic mRNA. *EMBO J* 17(7): 2107-2121.
88. Matsuoka K, Neuhaus J. 1999. Cis-elements of protein transport to the plant vacuoles. *J Exp B*ot 50:165–74.
89. May GD, Afza, R, Mason, HS, Wiecko, A, Novak, FJ and Arntzen, C J. 1995. Generation of transgenic banana (Musa acuminata) plants via Agrobacterium-mediated transformation. *BioTechnology* 13:486-492.
90. McKnight C. Crosswhite, FJ, Dossey JA, Kifer E, Swafford J O, Travers K J, and Cooney T J .1987. *The underachieving curriculum: Assessing U.S. school mathematics from an international perspective*. Champaign, IL: Stipes.
91. Miao Y, Ding Y, Sun QY, Xu ZF, Jiang L. 2008. Plant bioreactors for pharmaceuticals. *Biotechnol Genet Eng Rev* 25: 363-380.
92. Michaud, D, Vrain, TC, Gomord, V and Faye, L. 1998. Stability of recombinant proteins in plants. *Methods Biotechnol*. 3: 177–188.
93. Mishra S, Yadav DK, Tuli R. 2006. Ubiquitin fusion enhances cholera toxin B subunit expression in transgenic plants and the plant-expressed protein binds GM1 receptors more efficiently. *J Biotechnol* 127:95-108.

94. Mondal TK, Bhattacharya A, Ahuja PS, Chand PK .2001. Transgenic tea (Camellia sinensis [L.] O. Kuntze cv. Kangra Jat) plants obtained by Agrobacterium-mediated transformation of somatic embryos. *Plant Cell Rep*. 20:712-720
95. Murray C, Sutherland PW, Phung MM, Lester MT, Marshall RK, Christeller JT. 2002. Expression of biotin-binding proteins, avidin and streptavidin, in plant tissues using plant vacuolar targeting sequences. *Transgenic Res* 11(2): 199-214.
96. Nagaya S, Kawamura K, Shinmyo A, Kato K. 2010. The HSP terminator of Arabidopsis thaliana
97. Nandi S, Suzuki YA, Huang J, Yalda D, Pham P, Wua L, et al. 2002. Expression of human lactoferrin in transgenic rice grains for the application in infant formula. *Plant Sci* 163:713–22.
98. Nara Y, Kurata H, Seki M, Taira K. 2000. Glucocorticoid-induced expression of a foreign gene
99. by the GVG system in transformed tobacco BY-2 cells. *Biochem Eng* J 6185–91
100. Nuttall J, Vine N, Hadlington JL, Drake P, Frigerio L, Ma JK. 2002. ER-resident chaperone interactions with recombinant antibodies in transgenic plants. *Eur J Biochem* 269(24): 6042-6051.
101. Nykiforuk CL, Shen Y, Murray EW, Boothe JG, Busseuil D, Rhéaume E, et al. 2011. Expression and recovery of biologically active recombinant Apolipoprotein AIMilano from transgenic safflower (Carthamus tinctorius) seeds. *Plant Biotechnol*;9:250–63
102. Obembe OO, Popoola JO, Leelavathi S, Reddy SV. 2011. Advances in plant molecular farming. *Biotechnol Adv* 29(2): 210-222.
103. Parmenter DL, Boothe JG, van Rooijen GJ, Yeung EC, Moloney MM. 1995. Production of biologically active hirudin in plant seeds using oleosin partitioning. *Plant Mol Biol* 29(6): 1167-1180.

104. Peebles CAM, Gibson SI, Shanks JV, San KY. 2007.Characterization of an ethanol-inducible promoter system in Catharanthus roseus hairy roots. *Biotechnol Prog* 23: 1258–60.
105. Peng RH, Yao QH, Xiong AS, Cheng ZM, Li Y. 2006. Codon-modifications and an endoplasmic reticulum-targeting sequence additively enhance expression of an Aspergillus phytase gene in transgenic canola. *Plant Cell Rep* 25(2): 124-132.
106. Platis D, Labrou NE. 2008. Affinity chromatography for the purification of therapeutic proteins from transgenic maize using immobilized histamine. *J Sep Sci* 31:636–45.
107. Pogrebnyak N, Golovkin M, Andrianov V, Spitsin S, Smirnov Y, Egolf R, Koprowski H. 2005. Severe acute respiratory syndrome (SARS) S protein production in plants: development of recombinant vaccine. *Proc Natl Acad Sci U S A* 102(25): 9062-9067.
108. Pogue GP, Vojdani F, Palmer KE, Hiatt E, Hume S, Phelps J, et al. 2010. Production of pharmaceutical-grade recombinant aprotinin and a monoclonal antibody product using plant-based transient expression systems. *Plant Biotechnol J* 8:638–54.
109. Porta C, Lomonossoff GP. 2002. Viruses as vectors for the expression of foreign sequences in plants. *Biotechnol Genet Eng Rev* 19:245–91
110. Prakash CS. 1996. Edible vaccines and antibody producing plants. *Biotechnol Dev Moni*t 27:10–3.
111. Rajasekharan PE. 2006. Potential of plant derived pharmaceuticals. *AprilSci Tech Entrepreneur*: 4.
112. Ramessar K, Capell T, Twyman RM, Quemada H, Christou P. 2008b . Trace and traceability—a call for regulatory harmony. *Nat Biotechnol* 26:975–8

113. Rashid, H, Yokosi, S, Toriyama, K and Hinata K. 1996. Transgenic plant production mediated by Agrobacterium in Indica rice. *Plant Cell. Rep*. 15:727-730.
114. Reddy VS, Leelavathi S, Selvapandiyan A, Raman R, Giovanni F, Shukla V, et al. 2002. Analysis ofchloroplast transformated tobacco plants with cry1Ia5 under rice psbA transcriptional elements reveal high level expression of Bt toxin without imposing yield penalty and stable inheritance of transplastome. *Mol Breed* 9:259–69.
115. Reed DG, Nopo-Olazabal LH, Funk V, Woffenden BJ, Reidy MJ, Dolan MC, et al. 2005. Expression of functional hexahistidine-tagged ricin B in tobacco. *Plant Cell Rep* 24:15–24
116. Regnard GL, Halley-Stott RP, Tanzer FL, Hitzeroth II, Rybicki EP. 2010. High level protein expression in plants through the use of a novel autonomously replicating geminivirus shuttle vector. *Plant Biotechnol J* 8:38–46.
117. Richter LJ, Thanavala Y, Arntzen CJ, Mason HS. 2000. Production of hepatitis B surface antigen in transgenic plants for oral immunization. *Nat Biotechnol* 18:1167–7
118. Ritala A, Aspegren K, Kurten, U, SalmenkallioMarttila M, Mannonen L, Hannus,R, Kauppinen V, Teeri TH, Enari TM. 1994. Fertile transgenic barley by particle bombardment of immature embryos. *Plant Molecular Biology* 24:317-325.
119. Robinson DG, Oliviusson P, Hinz G. 2005. Protein sorting to the storage vacuoles of plants: a critical appraisal. *Traffic* 6(8): 615-625.
120. Rybicki EP.2010. Plant-made vaccines for humans and animals. *Plant Biotechnol J* 8: 620–37
121. Sandmann G, Romer S, Fraser PD. 2006. Understanding carotenoid metabolism as a necessity for genetic engineering of crop plants. *Metab Eng* 8(4): 291-302.

122. Scheller J, Conrad U. 2005. Plant-based material, protein and biodegradable plastic. *Curr Opin Plant Biol* 8(2): 188-196.
123. Scheller J, Guhrs KH, Grosse F, Conrad U. 2001. Production of spider silk proteins in tobacco and potato. *Nat Biotechnol* 19(6): 573-577.
124. Schillberg S, Zimmermann S, Voss A, Fischer R. 1999. Apoplastic and cytosolic expression of full-size antibodies and antibody fragments in *Nicotiana tabacum*. *Transgenic Res* 8(4): 255-263.
125. Schouten A, Roosien J, van Engelen FA, de Jong GA, Borst-Vrenssen AW, Zilverentant JF, Bosch D, Stiekema WJ, Gommers FJ, Schots A et al. 1996. The C-terminal KDEL sequence increases the expression level of a single-chain antibody designed to be targeted to both the cytosol and the secretory pathway in transgenic tobacco. *Plant Mol Biol* 30(4): 781-793.
126. Seki M, Shigemoto N, Komeda Y, Imamura J and Morikawa, H. 1991. Transgenic Arabidopsis thaliana plants obtained by particle bombardment mediated transformation. *Applied Microbiology and Biotechnology* 36:228-230.
127. Seon JH, Szarka JS and Moloney MM. 2002. A unique strategy for recovering recombinant proteins from molecular farming: affinity capture on engineered oilbodies. *J.Plant Biotechnology* 4 (3): 95-101.
128. Sharma AK, Sharma MK. 2009. Plants as bioreactors: Recent developments and emerging opportunities. *Biotechnol Adv* 27(6): 811-832.
129. Sharma MK, Singh NK, Jani D, Sisodia R, Thungapathra M, Gautam JK, et al. 2008. Expression of toxin co-regulated pilus subunit A (TCPA) of Vibrio cholerae and its immunogenic epitopes fused to choleratoxin B subunit in transgenic tomato (Solanum lycopersicum). *Plant Cell Rep* 27:307–18.

130. Shimao M. 2001. Biodegradation of plastics. *Curr Opin Biotechnol* 12(3): 242-247.
131. Shimoda N, Toyoda-Yamamoto A, Nagamine J, Usami S, Katayama M, Sakagami Y, Machida Y 1990. Control of expression of Agrobacterium vir genes by synergistic actions of phenolic signal molecules and monosaccharides *Proc. Natl. Acad. Sci*. USA 87:6684-6688
132. Somers DA, Rines, HW Gu W Kaeppler HF and Bushnell WR. 1992. Fertile, transgenic oat plants. *BioTechnology* 10:1589-1594.
133. Spencer PA, Towers GHN .1988. Specificity of signal compounds detected by Agrobacterium tumefaciens*. Phytochemistry* 27:2781-2785
134. Spok A, Twyman RM, Fischer R, Ma JK, Sparrow PA. 2008. Evolution of a regulatory framework for pharmaceuticals derived from genetically modified plants. *Trends Biotechnol* 26:506-517.
135. Spok A. 2007. Molecular farming on the rise—GMO regulators still walking a tightrope. *Trends Biotechnol* 25:74–82.
136. Sreeramanan S, Vinod B, Ranjetta P, Sasidharan S, Xavier R .2009. Chemotaxis movement and attachment of *Agrobacterium tumefaciens* to *Phalaenopsis violacea* orchid tissues: an assessment of early factors influencing the efficiency of gene transfer. *Trop Life Sci Res*. 20(1):39-49
137. Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll JA, Spatola L et al. 2000. High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nat Biotechnol* 18(3): 333-338.
138. Stoger E, Fink C, Pfosser M, Heberle-Bors E. 1995. Plant transformation by particle bombardment of embryogenic pollen. *Plant Cell Rep* 14(5): 273-278.

139. Streatfield SJ, Howard JA. 2003. Plant-based vaccines. *Int J Parasitol* 33(5-6): 479-493.
140. Sugio T, Matsuura H, Matsui T, Matsunaga M, Nosho T, Kanaya S, et al.2010. Effect of the sequence context of the AUG initiation codon on the rate of translation in dicotyledonous and monocotyledonous plant cells*. J Biosci Bioeng* 109: 170–3.
141. Teli, NP and Timko, MP. 2004. Recent developments in the use of transgenic plants for the production of human therapeutics and biopharmaceuticals. *Plant Cell Tissue and Organ Culture*, 79: 125-145.
142. Terpe K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*; 60:523–33.
143. Thomas BR, Van Deynze A and Bradford KJ. 2002. Production of Therapeutic proteins in plants. *Agricultural Biotechnology* in California Series, Publication 8078
144. Tiwari S, Verma PC, Singh PK, Tuli R. 2009. Plants as bioreactors for the production of vaccine antigens. *Biotechnol Adv* 27(4): 449-467.
145. Tomes, DT, Weissinger, AK, Ross M, Higgins, R, Drummond, BJ, Schaff, S, Malone-Schoneberg, J, Staebell M, Flynn P, Anderson J and Howard J. 1990. Transgenic tobacco plants and their progeny derived by microprojectile bombardment of tobacco leaves. *Plant Mol. Biol.* 14:261-268.
146. Tregoning JS, Nixon P, Kuroda H, Svab Z, Clare S, Bowe F, Fairweather N, Ytterberg J, van Wijk KJ, Dougan G et al. 2003. Expression of tetanus toxin Fragment C in tobacco chloroplasts. *Nucleic Acids Res* 31(4): 1174-1179.
147. Twyman RM, Schillberg S, Fischer R. 2005. Transgenic plants in the biopharmaceutical market. *Expert Opin Emerg Drugs* 10(1): 185-218.

148. Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R. 2003. Molecular farming in plants: host systems and expression technology. *Trends Biotechnol* 21(12): 570-578.
149. Valdez-Ortiz A, Rascon-Cruz Q, Medina-Godoy S, Sinagawa-Garcia SR, Valverde-Gonzalez ME, Paredes-Lopez O. 2005. One-step purification and structural characterization of a recombinant His-tag 11S globulin expressed in transgenic tobacco. *J Biotechnol* 115:413–23.
150. Van Rooijen GJH, Moloney MM. 1995. Plant seed oil-bodies as carriers for foreign proteins.*Bio-Technol* 13:72–7
151. Varsani A, Williamson AL, Stewart D, Rybicki EP.2006. Transient expression of human papillomavirus type 16L1 protein in Nicotiana benthamiana using an infectious tobamovirus vector. *Virus Res* 120:91–6.
152. Vasil V, Castillo AM, Fromm ME and Vasil IK. 1992. Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Biotechnol.* 10: 667-674.
153. Vezina LP, Faye L, Lerouge P, D'Aoust MA, Marquet-Blouin E, Burel C, et al. 2009.Transient coexpression for fast and high-yield production of antibodies with human-likeN-glycans in plants. *Plant Biotechnol J* 7:442–55.
154. Vidi PA, Kessler F, Brehelin C. 2007. Plastoglobules: a new address for targeting recombinant proteins in the chloroplast. *BMC Biotechnol* 7: 4
155. Vierstra RD. 1996. Proteolysis in plants: mechanisms and functions. *Plant Mol Biol* 32(1-2): 275-302.
156. Vitale A, Galili G. 2001. The endomembrane system and the problem of protein sorting. *Plant Physiol* 125: 115-118

157. Vitale A, Hinz G. 2005. Sorting of proteins to storage vacuoles: how many mechanisms? *Trends Plant Sci* 10(7): 316-323.
158. Vitale A, Pedrazzini E. 2005. Recombinant pharmaceuticals from plants: the plant endomembrane system as bioreactor. *Mol Interv* 5(4): 216-225.
159. Voinnet O, Rivas S, Mestre P, Baulcombe D. 2003. An enhanced transient expression system in
160. Wallroth M, Gerats AM, Rogers SG, Fraley RT, Horsch RB. 1986. Chromosomal localization of foreign genes in Petunia Hybrida. *Mol. Gen.Genet* 202, 6-15.
161. Wan Y and Lemaux PG. 1994. Generation of large numbers of independently transformed fertile barley plants. *Plant Physiol*. 104: 37-48.
162. Wandelt CI, Khan MR, Craig S, Schroeder HE, Spencer D, Higgins TJ. 1992. Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. *Plant J* 2(2): 181-192.
163. Wilken LR, Nikolov ZL. 2012. Recovery and purification of plant-made recombinant proteins. *Biotechnol Adv* 30(2): 419-433.
164. Witte CP, Noel LD, Gielbert J, Parker JE, Romeis T. 2004. Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. *Plant Mol Biol* 55:135–47.
165. Wu G, Truksa M, Datla N, Vrinten P, Bauer J, Zank T, Cirpus P, Heinz E, Qiu X. 2005. Stepwise engineering to produce high yields of very long-chain polyunsaturated fatty acids in plants. *Nat Biotechnol* 23(8): 1013-1017.
166. Wu H, McCormac AC, Elliott MC, Chen DF. 1998. Agrobacterium-mediated stable transformation of suspension cultures of barley (Hordeum vulgare L.). *Plant Cell Tissue Organ Cult* 54: 161–167

167. Yang D,Wu L, Hwang YS, Chen L, Huang N. 2001. Expression of the REB transcriptional activator in rice grains improves the yield of recombinant proteins whose genes are controlled by a Reb-responsive promoter. *Proc Natl Acad Sci USA* 98:11438–43
168. Yang J, Barr LA, Fahnestock SR, Liu ZB. 2005. High yield recombinant silk-like protein production in transgenic plants through protein targeting. *Transgenic Res* 14(3): 313-324.
169. Yoshida, K, Matsui, T. and Shinmyo, A. 2004. The plant vesicular transporengineering for production of useful recombinant proteins. *J. Mol. Catal. B: Enzym*. 28:167–171.
170. Youm JW, Jeon JH, Kim H, Kim YH, Ko K, Joung H, Kim H. 2008. Transgenic tomatoes expressing human beta-amyloid for use as a vaccine against Alzheimer's disease. *Biotechnol Lett* 30(10): 1839-1845.
171. Ytterberg AJ, Peltier JB, van Wijk KJ. 2006. Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiol* 140(3): 984-997.
172. Zhong XB, Hans de, Jong J, Zabel P .1996. Preparation of tomato meiotic pachytene and mitotic metaphase chromosomes suitable for fluorescence in situ hybridization (FISH). *Chrom Res* 4: 24-28.
173. Zimmermann MB, Hurrell RF. 2002. Improving iron, zinc and vitamin A nutrition through plant biotechnology. *Curr Opin Biotechnol* 13(2): 142-145.

Chapter 5
Plant Transgenics: Genetic Engineering Approch to Devlop Biotic Stress Resistance Plants

Chakravarthi. M, Vinoth. S and Chandrashekara. K.N

Abstract

Diseases caused by microorganisms are currently the major factor limiting crop production worldwide. In addition to negative effects on yield, diseases can also influence the post-harvest quality of food. Due to high cost, efficacy and environmental concerns, much research is presently aimed at expression of transgenes that can confer significant levels of disease resistance. The genetic manipulation of plants has been going on since the dawn of agriculture, but until recently this has required the tedious process of cross-breeding varieties. Genetic engineering promises to speed the process and broaden the scope of what can be done. To date, most interest has been focused on developing virus resistant transgenic plants, but through biotechnology to confer resistance to fungi, bacteria, or nematodes has also been gaining great consideration. Although recent introductions of plant products for control of insect pests have been highly successful, transgenic plants exhibiting resistance to fungal or bacterial diseases have yet to reach the marketplace. This book chapter mainly focusses on the novel strategies that are being manipulated for the development of disease resistant transgenic plants.

INTRODUCTION

The breakthroughs in science that permitted genes and thus heredity, to be identified and manipulated as molecules ushered in the biotechnology era, which is now more than two decades old. The new tools of biotechnology are changing the way scientists can address problems in the life sciences; agriculture is one area facing major changes as a result of this new technology. The unanticipated rapid rate at which discoveries and their applications in biotechnology have unfolded has stressed the capacity of society more specifically, our agricultural research and educational institutions to absorb and adjust to change. We are challenged by pressing decisions, opportunities and problems that we face now and will continue to face in the future. Competition from abroad impels us to devise and use new technologies that can improve the efficiency and quality of Indian agricultural production. Securing a better life for our citizens where each one of us, can lead lives of dignity and fulfillment therefore merits undivided attention in our development strategy. A natural corollary to this is the attainment of the goal of food security for all. The per capita availability of land has also been shrinking due to population increase and this has been compounded by increase of wasteland. There is also spread of urbanization and the growing demand for more land. We are left with a situation where we have to produce more from the limited land available to ensure food security. But in doing so, we must always keep in mind that any food production and consumption policy must safeguard that the integrity of natural eco-systems is not compromised.

THE POWER OF BIOTECHNOLOGY

The power of biotechnology is no longer fantasy. Biotechnology: the use of technologies based on living systems to develop commercial processes and products now includes the techniques of recombinant

DNA, gene transfer, plant regeneration, cell culture, monoclonal antibodies and bioprocess engineering. Using these techniques, we have begun to transform ideas into practical applications. For instance, scientists have learned to genetically alter certain crops to increase their tolerance to certain bacteria, viruses, fungi etc. Biotechnology offers new ideas and techniques applicable to agriculture. It offers tools to develop a better understanding of living systems, of our environment and of ourselves. Yet continued advances will take a serious commitment of talent and funds. Biotechnology offers tremendous potential for improving crop production. It can provide scientists with new approaches to develop crop varieties resistance to diseases and reduce the need for expensive agricultural chemicals.

STRATEGIES FOR COMPETITIVENESS

It is important to develop a national strategy for biotechnology in agriculture because biotechnology offers opportunities for increased sustainability, profit ability and international competitiveness in agriculture. Such a strategy should address improving the full spectrum of activities, from the quality and direction of research to the realization of the benefits of this research in agricultural production.

The potential benefits of biotechnology will not be realized without a continued commitment to basic research. Six research areas on merit emphasis:

1. **Gene identification:** Locating and identifying agriculturally important genes and creating chromosomemaps for elite varieties.
2. **Gene regulation:** Understanding the mechanisms of regulation and expression of these genes and refiningthe methods by which they may be genetically engineered.
3. **Structure and function of gene products:** Understanding the

structure and function of gene productsin metabolism and the development of agriculturally important traits.

4. **Cellular techniques:** Developing and refining techniques for cell culture, cell fusion, regeneration ofplants and other manipulations of plant cells and embryos.
5. **Development in organisms and communities:** Understanding the complex physiological and geneticinteractions and associations that occur within an organism and between organisms.
6. **Environmental considerations:** Understanding the behavior and effect of genetically engineeredorganisms in the environment.

TOOLS OF GENETIC ENGINEERING IN PLANTS

Transfer and expression of foreign genes in plant cells, now routine practice in several laboratories around the world, has become a major tool to carry out gene expression studies and to obtain plant varieties of potential agricultural interest. The capacity to introduce and express diverse foreign genes in plants, first described for tobacco in De Block (1984), has been extended to many species. Transgenic crops such as tomato, papaya, cotton, maize, soybean etc., are now available for human consumption and by complementing traditional methods of crop improvement (and thus becoming an integral part of agriculture), they will have a profound impact on food production, economic development and on the development of a sustainable agricultural system during the 21st century. Although the capacity to introduce and manipulate specific gene expression in plants provides a powerful tool for fundamental research, much of the support for plant transformation research has been provided because of the generation of plants with useful and rapidly discernible phenotypes which are unachievable by conventional

plant breeding i.e., resistance to viruses, insects, herbicides, or post-harvest deterioration (Nelson *et al*., 1988; Staskawicz *et al*., 1995). In this chapter the technical aspects of the state of the art in plant engineering are described. It also identifies technical problems remaining in the development of systems of plant transformation applicable to crop improvement.

DNA Delivery Systems

Agrobacterium tumefaciens

This bacterium is a natural transformer of somatic host cells of plants into tumorous crown gall cells. Its ability to transform cells with a piece of DNA was exploited by plant biologists, and now *Agrobacterium* plays a prominent role in transgenesis of plants. This natural gene transfer system is highly efficient, frequently yielding transformants containing single copies of the transferred DNA which have a relatively uncomplicated integration pattern compared with other transformation procedures. Its utility was developed from the understanding of the molecular basis of the crown gall disease, namely, the transfer of DNA from the bacterium to the plant nuclear genome during the tumor formation process. Only a small discrete portion of the ca. 200 kbp tumor-inducing plasmid (Ti) existing in the bacterium is transferred to the plant genome. The transferred DNA, now familiarly referred to as T-DNA, is surrounded by two 25-bp imperfect direct repeats and contains oncogenes encoding enzymes for the synthesis of the plant growth regulators auxin and cytokinin and for the synthesis of novel amino acid derivatives called opines. The DNA transfer is mediated by a set of bacterial proteins encoded by genes (*vir* genes) existing in the Ti-plasmid, which become induced by phenolics compounds released upon wounding of the plant tissue. The key aspect in regard to gene transfer is that none of the T-DNA genes are involved in the transfer process and

therefore, any or all of these genes can be removed, mutated or replaced by other genes, and the T-DNA region can still be transferred to the plant genome.

Direct Gene Transfer

For some time there was good reason to believe that *Agrobacterium tumefaciens* was the vector system with the capacity for gene transfer to any plant species and variety. As this was not the case, numerous alternative approaches of 'direct gene transfer' have been tested. Most methods of direct gene transfer, such as the introduction of DNA via electroporation, (Riggs and Bates, 1986; Christou *et al*., 1987; Shimamoto *et al*., 1989). PEG-mediated DNA uptake, (Hayashimoto *et al*., 1990; Torres *et al*., 1997), protoplast fusion with liposomes containing DNA (Caboche, 1990), biolistics (Christou, 1992) or microinjection (Schnorf *et al*., 1991), require the regeneration of plants from protoplasts. The recalcitrance of many plant species for efficient regeneration from protoplasts, elaborate protocols and prolonged tissue culture phases, are a disadvantage. Other methods for direct gene transfer in which DNA is introduced directly into tissue or whole plants (Christou *et al.*, 1991; D'halluin *et al*., 1992; Chowrira *et al*., 1995; Klein *et al*., 1987; Barcelo *et al*., 1994) do not require protoplasts. Biolistics, or acceleration of heavy microparticles coated with DNA, has been developed into a technique that carries genes into virtually every type of cell and tissue. Without too much manual effort, this approach has advantages such as easy handling, regeneration of multiple transformants in one shot and utilization of a broad spectrum of target cells, i.e., pollen, cultured cells, meristematic cells, etc. Using this technique, a number of transgenic crops have been produced. Electroporation is one of several standard techniques for routine and efficient transformation of plants from protoplasts (Riggs and Bates,

1986; Fromm *et al*., 1987; Fromm *et al*., 1986)**.** This technique refers to the process of applying a high-intensity electric field to reversibly permeabilize bilipid membranes and it may be applicable to all cell types. Discharge of a capacitor across cell populations leads to transient openings in the plasmalemma which facilitates entry of DNA molecules into cells if the DNA is in direct contact with the membrane. Transgenic plants recovered using this technique contain from one to few copies of the transfected DNA, which is generally inherited in a Mendelian fashion.

The Selection and Analysis of Transformants

Using either *Agrobacterium* or direct gene transfer systems, it is now possible to introduce DNA into virtually any regenerable plant cell type. However, only a minor fraction of the treated cells become transgenic while the majority of the cells remain untransformed. It is therefore essential to detect or select transformed cells among a large excess of untransformed cells, and to establish regeneration conditions allowing recovery of intact plants derived from single transformed cells.

Selectable genes

Selectable marker genes are essential for the introduction of agronomically important genes into important crop plants. The agronomic gene(s) of interest are invariably co-introduced with selectable marker genes and only cells that contain and express the selectable marker gene will survive the selective pressure imposed in the laboratory. Plants regenerated from the surviving cells will contain the selectable marker joined to the agronomic gene of interest. The selection of transgenic plant cells has traditionally been accomplished by the introduction of an antibiotic or herbicide-resistant gene, enabling the transgenic cells to be selected on media containing the corresponding toxic compound. The antibiotics and herbicides selective agents are used only in the laboratory in the initial stages of the genetic modification

process to select individual cells containing genes coding for agronomic traits of interest. The selective agents are not applied after the regeneration of whole plants from those cells nor during the subsequent growth of the crop in the field. Therefore, these plants and all subsequent plants and its products will neither have been exposed to, nor contains the selective agent. By far, the most widely used selectable gene is the neomycin phosphotransferase II (*NPTII*) gene (Fraley, 1986) which confers resistance to the aminoglycoside antibiotics kanamycin, neomycin, paromomycin and G-418 (Bevan *et al*., 1983; Guerche *et al*., 1987). Another widely used selectable marker is the *hptII* gene. The HPTII gene derived from *E. coli* is an aminocyclitol antibiotic that interferes with protein synthesis. The bacterial hpt gene was modified for expression in plants (Waldron *et al.,* 1985) and has then been widely used as selectable marker in plant transformation. Hygromycin is generally more toxic to cells than kanamycin and quick as well as effective in killing of sensitive cells. Transformants are selected by applying hygromycin concentrations ranging from 20-200mg/l. It has been employed to transform diverse plant species, especially grasses in which *nptII* is ineffective (Miki and McHugh, 2004).

A number of other selective systems have been developed based on resistance to bleomycin (Hille *et al*., 1986), bromoxynil, chloramphenicol (Fraley, 1983).

Reporter genes

Reporter genes are 'scoreable' markers which are useful for screening and labeling of transformed cells as well as for the investigation of transcriptional regulation of gene expression. Furthermore, reporter genes provide valuable tools to identify genetic modifications. They do not facilitate survival of transformed cells under particular laboratory conditions but rather, they identify or tag

transformed cells. They are particularly important where the genetically modified plants cannot be regenerated from single cells and direct selection is not feasible or effective. They can also be important in quantifying both transformation efficiency and gene expression in transformants. The reporter gene should show low background activity in plants, should not have any detrimental effects on plant metabolism and should come with an assay system that is quantitative, sensitive, versatile, simple to carry out and inexpensive. A main use of reporter genes is promoter characterization. They are transcriptionally fused to the promoter of interest and assayed to determine the expression conferred by the promoter. They are also used in gene silencing approaches, in studying transposon activity, as markers for specific cellular compartments and in selection of transformed cells and tissues (Rosellini, 2012). Arrays of reporter genes have been reported so far of which the most widely used include uidA (GUS), lacZ (β-galactosidase), GFP (Green fluorescent protein) and Luc (Luciferase). In addition, several anthocyanin pigmentation genes are also used as RGs due to their stability and easy detection.The gene encoding for the enzyme-glucuronidase, *GUS*, has been developed as a reporter system for the transformation of plants (Jefferson *et al*., 1986;Vancanney *et al*., 1990).The ß- glucuronidase enzyme is a hydrolase that catalyzes the cleavage of a wide variety of ß -glucuronides, many of which are available commercially as spectrophotometric, fluorometric and histochemical substrates. There are several useful features of *GUS* which make it a superior reporter gene for plant studies. Firstly, many plants assayed to date lack detectable *GUS* activity, providing a null background in which to assay chimaeric gene expression. Secondly, glucuronidase is easily, sensitively and cheaply assayed both *in vitro* and *in situ* in gels and is robust enough to withstand fixation, enabling histochemical localization in cells and tissue sections. Thirdly, the enzyme tolerates

large amino-terminal additions, enabling the construction of translational fusions. The gene encoding firefly luciferase has proven to be highly effective as a reporter because the assay of enzyme activity is extremely sensitive, rapid, easy to perform and relatively inexpensive. Light production by luciferase has the highest quantum efficiency known of any chemiluminescent reaction. Additionally, luciferase is a monomeric protein that does not require posttranslational processing for enzymatic activity (De Wet *et al*., 1985).

The use of green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* to label plant cells has become an important reporter molecule for monitoring gene expression *in vivo*, *in situ* and in real time. GFP emits green light when excited with UV light. Unlike other reporters, GFP does not require any other proteins, substrates or cofactors. GFP is stable, species-independent and can be monitored noninvasively in living cells. It allows direct imaging of the fluorescent gene product in living cells without the need for prolonged and lethal histochemical staining procedures. In addition, GFP expression can be scored easily using a long-wave UV lamp if high levels of fluorescence intensity can be maintained in transformed plants. Another advantage of GFP is that it is relatively small (26 kDa) and can tolerate both N- and C-terminal protein fusions, lending itself to studies of protein localization and intracellular protein trafficking (Kaether and Gerdes, 1995). It has been reported that high levels of GFP expression could be toxic to plant growth and development (Rouwendal *et al*., 1997). Solution to this problem comes from the utilization of GFP mutant genes. Among the various GFP mutations, the S65T (replacement of the serine in position 65 with a threonine) is one of the brightest chromophores characterized by its faster formation and greater resistance to photo bleaching than wild-type GFP photo bleaching. Furthermore, this mutant is characterized by having a single excitation peak ideal for fluorescin isothiocyanate filter

sets (Heim *et al*., 1995) and also by its harmless action to the plant cell (Niwa *et al*., 1999).

Alternatives to antibiotic resistance markers

Alternative selectable markers for plants fall into two categories. Some markers confer resistance to chemicals other than antibiotics that kill plant cells such as herbicides and lethal concentrations of the amino acids lysine and threonine. The enzyme that confers resistance to high concentrations of lysine and threonine can interfere with amino acid biosynthesis and if expressed at high levels cause abnormal plant development. The relevant genes are therefore not suitable as marker systems. Other alternative marker systems rely on the growth of plant cells in the presence of unusual nutrients, including cytokinin, glucuronides, xylose or mannose, which will not allow non-transformed plant cells to grow. For example, plant cells usually do not use mannose as a source of sugar. The delivery of a gene allowing mannose to be metabolised in plant cells and the subsequent cultivation of those cells in a medium containing mannose as the sole source of sugar would allow only those cells which have taken up the gene to grow. When these systems, which are still in their development phase, work reliably on a large scale in a wide range of different environments risk assessments will have to be conducted to assess the potential ecological impacts of plants that can grow on a new substrate, the impact on the overall plant metabolism and the consequences on human or animal diet from increased levels of metabolites in these crops that might not be present in the conventional counterpart.

Removal of Markers

It is not possible to remove marker genes once they are integrated into a plant genome unless a particular mechanism for

removal is incorporated along with the marker gene and the gene of interest at the time of the transformation. As was mentioned above, it is possible to avoid introducing into plant cells antibiotic resistant marker genes which are only used for the assembly and amplification of the DNA constructs in bacteria, and therefore are not necessary during the plant step of the transformation procedure. The removal prior to commercialisation of marker genes which are driven by plant promoters and are used for selection of plant cells has become the aim of both consumers and industry. Extensive research with this aim is being carried out both by industry and academic institutions. Among the technologies being assessed are:

1. The use of meganucleases (e.g.: Cre/lox system). These are enzymes which can specifically recognise long DNA sequences. These recognition sequences are introduced on both sides of the antibiotic resistant marker gene to be introduced into the plant cell. Once the transformed cells have been selected on the corresponding antibiotic, the meganuclease is introduced into the plant cell, and will allow the excision of the antibiotic resistant marker gene. This technology has proven to be very efficient in certain plants, but difficult to handle in others possibly because the meganuclease recognises sites in the plant genome itself.
2. The presence of homologous DNA sequences on both side of the antibiotic resistant marker gene may allow for random recombination and elimination of the gene. This process of homologous recombination occurs at low frequency and may be plant specific.
3. It is possible to introduce the trait of interest and the antibiotic resistant marker on different DNA constructs. Following transformation, each molecule integrates on a different

chromosome. In this case it is possible to segregate the trait of interest from the marker gene at the next generation. Frequencies of integration on separate chromosomes can be quite low when compared to integration at the same locus.

Modification of regulatory elements of marker genes

The concern about antibiotic resistance marker genes is predominantly about their transfer and expression in bacterial cells and a technology which would prevent such expression might have to be considered. Already the genes used for selection in plants are controlled by plant promoter sequences which render them unlikely to be sufficiently expressed in bacteria. The introduction of an intron sequence in the marker gene would restrict its expression to plant cells and definitively prevent any expression in bacteria. Introns are sequences of DNA that naturally interrupt the coding sequences of animal and plant cells. These are equipped with mechanisms allowing their removal during the transcription process while bacteria are not equipped to do this and therefore would be unable to read a gene containing introns.

Antibiotic resistance marker genes such as the *NPTII* gene which provides resistance towards kanamycin are very well researched in all the relevant aspects such as their functioning, biochemical properties and prevalence in the bacterial community. Their safety has been well examined and assessed. Under these conditions it is likely that achieving the same level of confidence as has been established for *NPTII* with another selection system may be long and difficult. Any remaining concerns attached to such genes could be removed by the addition of an intron.

More 'Friendly' Selectable Markers: the Positive Selection Method

In contrast to the traditional selection where the transgenic cells acquire the ability to survive on selective media while the non-transgenic cells are killed (negative selection), the positive selection method, first developed by Joersbo and Okkels (1996) favors regeneration and growth of the transgenic cells while the non-transgenic cells are starved but not killed. The positive selection method exploits the fact that cytokinin must be added to plant explants in order to obtain optimal shoot regeneration rates. By adding cytokinin as an inactive glucuronide derivate, cells which have acquired the *GUS* gene by transformation are able to convert the cytokinin glucuronide to active cytokinin while untransformed cells are arrested in development. In this system, *GUS* serves the dual purpose of being both a selectable and screenable marker gene. Another interesting system of positive selection uses the xylose isomerase gene from *Thermoanaerobacterium thermosulforogenas* as a selectable gene, which expression allows effectiveselection of transgenic plan cells using D-xylose as the selection agent (Haldrup *et al*., 1998). The transformation frequencies obtained by positive selection appear to be higher than using the negative selection method. This could be related to the fact that during negative selection the majority of the cells in the explants die. Such dying cells may release toxic substances which in turn may impair regeneration of the transformed cells. In addition, dying cells may form a barrier between the medium and the transgenic cells preventing uptake of essential nutrients.

BIOTIC STRESS

There is an increasing demand by consumers for fruits and vegetables free of pesticide and other residues, but cultivation without their use is only partially possible by using suitable resistant genotypes in a suitable environment. Plants have developed several natural defence

strategies to protect themselves against attack of pathogen and pest diseases (Hammond-Kosack and Jones, 1996). Concerning pathogen infection, strategies fall mainly into two groups:

1. ***Specific mechanisms*** responsible for pathogen recognition and control by a specific resistance againsta specific pathogen, with a hypersensitive disease resistance response (HR), hampering the diffusion of pathogen to healthy tissues by formation of necrotic lesion;
2. ***General mechanisms*** that confer resistance to a broad range of pathogens, occurring either in resistantor susceptible plants, but are able to control the pathogen. The synthesis of antimicrobial metabolites, lytic enzymes, pathogenesis-related proteins and other compounds strengthening the cell wall are involved. The resistance normally depends on the early response of the plant to pathogen attack, which lead to a rapid accumulation of reactive oxygen species (ROS) namely oxidative burst (Lamb and Dixon, 1997), with an accumulation of H_2O_2 which functions as a diffusable signal for the induction of cellular protectant genes (Delledonne *et al.,* 1999); nitric oxide co-operates in the induction of hypersensitive cell death. In some cases the plants react to pathogens by accumulating high levels of specific proteins which are toxic or inhibitory against both pathogens and pests (Broekaert *et al.,* 1995) such as RIP proteins, effective against insects and fungi; while other proteins seem to be more specific. Over expressing the genes by genetic engineering or induced mutation (Barbieri *et al.,* 1997; Maddaloni *et al.,* 1999) in plant cells under toxin or culture filtrate pressure are the two main strategies currently used to produce resistant plants. More research is needed to discover new molecular signals and the efficacy of the promoters of some genes involved in the defense; maybe the reinforcement of the promoters is sufficient to enhance plant resistance.

VIRUS RESISTANCE

Plant viruses reduce both the quantity and quality of crop yields by direct damage to plants, increasing sensitivity to adverse climatic conditions and to the direct pathogens. They cause trillions of Rupees of losses every year to crops worldwide, second only to the impact of fungal diseases (Waterworth and Hadidi, 1998). In several fruit crops virus diseases represent a particular problem, for example in grape with GCMV and GFLV, in *Prunus* spp. with Sharka and in some tropical species, such as papaya, with PRSV (Gonsalves, 1998). At present viral diseases are controlled in a number of ways including: planting virus-free plants, maintaining plant health, controlling plant pathogens which can be virus vectors and by cross protection (Alrefai and Korban, 1995). However, these techniques provide only limited protection from viral attack. Whilst, in the case of fungi, chemical defenses are available, such remedies are either not effective in the case of viruses or can make the impact of the virus even worse. The preventive use of resistant genotypes is thus essential (Khetarpal *et al.,* 1998). Two types transgenic resistance are available:

1. Pathogen-derived resistance (PDR) (used most at present)
2. Resistance induced by sequences of alien DNA.

PDR is conferred to the plants by genes from the virus itself, cloned and transferred to the host genome (Sanford and Johnston, 1985). PDR is developed when the viral gene products or virus-related sequences in the plant genome interferes with the virus infection cycle. The mechanisms which confer PDR are not yet well understood, varying with the nature of the gene used (Carr and Zaitlin, 1993; Fitchen and Beachy, 1993; Baulcombe, 1994; Kaniewski and Lawson, 1998; Yie and Tien, 1998; Martelli *et al.,* 1999; Smyth, 1999). Transgenic plants for the virus coat protein gene provide the most common strategy for gene transfer. The other strategies include antisense nucleic acids, satellite

sequences, defective interfering molecules and non-structural genes (replicase, protease, and movement proteins), antibodies, and interferon-related proteins (Gadani *et al.,* 1990; Baulcombe, 1994; Grumet, 1994; Kaniewski and Lawson, 1998; Wilson, 1993). Although a large number of crop plants has been successfully engineered using such strategies, for fruit crops only the coat protein strategy has been applied to confer PDR to potyvirus, nepovirus and closterovirus groups.

Studies demonstrate that this strategy is very promising, although in papaya Tennant *et al.* (1994) reported that CP-PRV was effective in protecting from some virus isolates but not from others. Studies by Singh *et al.* (1997) demonstrated that in tobacco, as a model plant, transgenic plants expressing a defective replicase gene of cucumber mosaic virus (CMV-FNY), acquired resistance to various banana isolates of CMV, suggesting this approach is worth further development. In most cases resistance has been successfully tested *in vivo* or indirectly by testing the accumulation of coat protein by ELISA or Western blot analysis or *gus* gene expression in the transgenic tissues. Examples of the resistance induced by sequences of alienDNA are not yet available but it should be possible to obtain them since in some species, such as *Citrus* spp., resistance to CTV is present in *Poncirus trifoliata* and is known to be controlled by a dominant gene at the Ctr locus. Developing transgenic fruits for virus resistance may lead to possible risks.

These include:

- Trans capsidation, when nucleic acids of a virus are covered by the coat protein belonging to another virus expressed by the transgenic plant (Farinelli *et al.,* 1992; Greene and Allison, 1994; Robinson *etal.,* 1999; Buzkan *et al.,* 2000). This problem is; however, already frequent in nature, with virusmultiple infections (Creamer and Falk,

1990; Hobbs and McLaughlin, 1990; Bourdin and Lecoq, 1991; Buzkan *et al.,* 2000);

- Recombination of nucleic acid expressed by the transgenic plants with nucleic acids of the virus occurring in transgenic plants, producing new more virulent viruses (Rybicki, 1994; Dolja *et al.,* 1994; Miller *et al.,* 1997; Aziz and Tepfer, 1999; Smith *et al.,* 2000). This problem is also very common in nature and together with mutations, is responsible for much viral evolution (Roossinck, 1997). According to the studies of Miller *et al.,* (1997), Jacquemond and Tepfer (1998), and other scientists, transgenic plants expressing viral sequences do not represent a source of risk greater than those already present in nature; Genetic depletion caused by abandoning susceptible varieties in favour of transgenic ones. This is a false problem since resistance can be conferred to susceptible varieties by biotechnologies;
- Compatible wild species which could become resistant following pollination with transgenic pollen produced by the transgenic crops. This is not usually a problem in areas where fruit/vegetable crops are cultivated because there are no wild relatives, except for the area of origin of the crop in question. Several fruit crops have been transformed with virus coat proteins; some of them showed resistance in field conditions, others have not been tested yet. An indirect strategy to fight viruses is to make plants resistant to their vectors. Yang *et al.* (2000) for example, have tried to make plants resistant to aphids, which are the vectors of grapefruit tristeza virus.

RNAi TECHNOLOGY

A decade has passed since the discovery that double-stranded RNA molecules (dsRNA) can trigger silencing of homologous genes, and it is now clear that RNA-mediated gene silencing is a widely conserved

cellular mechanism in eukaryotic organisms. The definition says RNA induced silencing of specific miRNA / gene is called RNA interface also known as gene silencing. In plants it is called as Post transcriptional silencing and in fungi it is known Quelling. RNAi mechanism evolved to provide immunity against various viruses and transposons. RNA-mediated gene silencing can be categorized into two partially overlapping pathways; the RNA interference (RNAi) pathway and the micro-RNA (miRNA) pathway. RNAi is triggered by either endogenous or exogenous dsRNA, and silences endogenous genes carrying homologous sequences at both the transcriptional and posttranscriptional levels. In contrast, the miRNA pathway is triggered by mRNAs transcribed from a class of non-coding genes. These mRNAs form hairpin-like structures, creating double-stranded regions in a molecule (pre-miRNA). In either pathway, dsRNA molecules are processed by Dicer RNase III proteins into small RNAs, which are then loaded into silencing complexes. In the RNAi pathway, small RNAs are called short interfering RNAs (siRNAs) and are loaded into RNA induced silencing complexes (RISC) for post-transcriptional silencing, or RNA-induced initiation of transcriptional gene silencing (RITS) complexes for transcriptional silencing. In contrast, miRNAs (small RNAs in the miRNA pathway) are loaded into miRNA ribonucleoparticles (miRNPs) for a review of silencing complexes). dsRNA binding motif (dsRBM) proteins, such as R2D2 and Loquacious, help small RNAs to be loaded properly into silencing complexes. Using the small RNA as a guide, silencing complexes find target mRNAs and cleave them (in the case of RISC) or block their translation (in the case of miRNP). RITS is involved in transcriptional silencing by inducing histone modifications. Argonaute family proteins are the main components of silencing complexes, mediating target recognition and silencing. The RNAi pathway and miRNA pathway are essentially parallel, using related but distinct proteins at each step.

RNAi is sequence-specific gene silencing at the post-transcription level, induced by double stranded RNA (dsRNA). The mediators of sequence-specific mRNA degradation are 21–23 nucleotides long, small interfering RNAs (siRNA) generated from longer dsRNAs by ribonuclease III cleavage activity (Fire *et al.,* 1998; Cogoni and Macino, 2000; Hannon, 2002; Agrawal *et al.,* 2003). The presence of RNAi machinery in insects has already been confirmed in earlier studies by introducing long dsRNA molecules into insects by means of using insect cell line (Zhang and Shi, 2002) or through injection into insect body (Soreq *et al.,* 1994) and also via forced feeding (Dong and Friedrich, 2005). That siRNAs are powerful agents for gene silencing, even at low concentrations, and mediate post-transcriptional degradation. Systemic RNAi was first described in plants as spread of post transcriptional gene silencing. The first animal in which RNAi was shown to work systemically was *C. elegans*, where it has been thoroughly investigated.

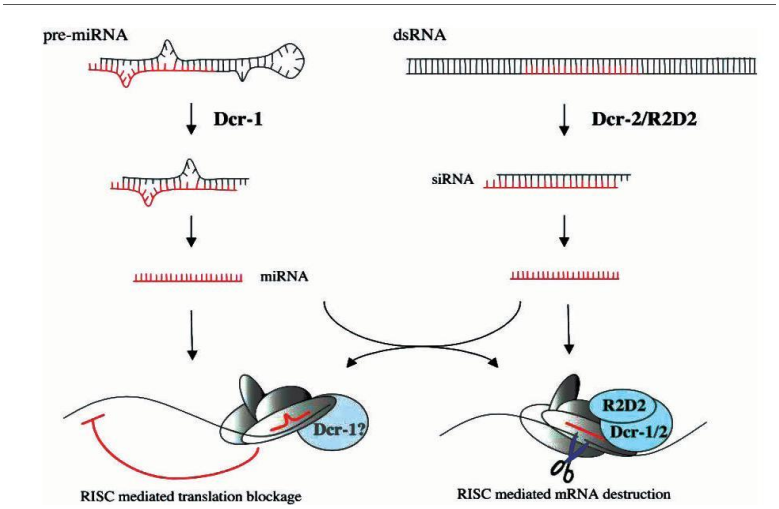


Fig. 1: Model for RNA Silencing in Drosophila: an ordered biochemical pathway, miRNAs (left panel)and siRNAs (right panel) are

processed from double-stranded precursor molecules by Dcr-1 and Dcr-2, respectively, and stay attached to Dicer-containing complexes, which assemble into RISC. The degree of complementarity between the RNA silencing molecule (in red) and its cognate target determines the fate of the mRNA: blocked translation or immediate destruction.

Although current understanding of RNAi activity cannot provide us with a precise prediction of potency to each individual siRNA. Once the bioinformatics part is done, candidate siRNA can then be synthesized and tested in cell culture systems for knockdown efficiency. The off-target effect can also be checked with a microarray assay. The final goal of this stage is to identify several siRNAs that show high knockdown efficiency and minimal off-target effect at nanomolar or lower concentrations.

Systematic Application of RNAi Technology

A growing interest is to exploit RNAi technology for systematic biology and functional genomics research to knockdown gene expression in whole-genome or whole-pathway scales. In Conclusions RNAi technology is a very timely invention in the era of post genome to serve as one of the most powerful tools for reverse genetics, functional genomics, and systematic biology. Although we should keep in mind that continuous improvements and modifications are necessary to make RNAi technology more potent, it already revolutionizes our way of doing biology and medicine. The explosion of RNAi technology actually benefited from the previous development of antisense technology. The discovery of the RNAi/miRNA pathway opens the door to RNAi technology, and further characterization of this pathway really facilitates the development of RNAi technology step by step.

FUNGAL RESISTANCE

Among diseases, fungi are the main cause of yield loss in fruit crops. They are controlled by several traditional techniques including quarantine, sanitation, breeding and clonal selection of resistant varieties and application of fungicides. However, resistant cultivars, with the onset of new strains of virulent pathogens, tend to become susceptible over time. In addition, the unrestrained use of fungicides, as well as increasing production costs and degrading the environment, induce new forms of resistance within pathogens, forcing the development of new pesticides. These problems have encouraged the search for biotechnological solutions to combating fungal disease. At present research is focused on identifying the genes involved in resistance, both those encoding for enzymes involved in the biosynthesis of toxic compounds for fungi and those encoding toxin proteins which directly inhibit fungal growth (Cornelissen and Melchers, 1993; Terras *et al.,* 1998) with the aim of introducing them in susceptible plants or substituting their inefficient antifungal gene promoters with more efficient ones. Several proteins have been reported with antifungal activity; they were classified into at least 11 classes named pathogenesis-related proteins (PRs). Some of them also showed antiviral and antibacterial activities.

Some defense-related genes encode enzymes involved in:

1. Phenyl propanoid metabolism;
2. Hydrolytic enzymes, such as chitinases and ß-1, 3–glucanases;
3. Hydroxyproline-rich glycoproteins (cell wall proteins);
4. Inhibitors of fungal enzymes, such as PGIP.

Plant ß-1, 3-Glucanases (PR-2) and chitinases (PR-3) represent potential antifungal hydrolases which act synergistically to inhibit fungal growth *in vitro* (Mauch *et al.,* 1988). In addition, ß-1, 3-Glucanases release glycosidic fragments from both the pathogen and the host cell

walls which could act as signals in the elicitation of host defences (Keen and Yoshikawa, 1983; Hahn *et al.,* 1989; Takeuchi *et al.,* 1990). Many of the genes induced by plant disease- resistance responses encode proteins with direct antifungal activity (AFPs) *in vitro* (Lamb *et al.,* 1992; Terras *et al.,* 1998). Identification of such anti-fungal proteins was isolated from plants and from the fungus itself such as *Tricoderma harsianum* (Neuhaus *et al.,* 1991; Mikkelsen *et al.,* 1992; Melchers *et al.,* 1993) and from humans. They include: Defensins, small cysteine-rich peptides, 2S albumins, chitin-binding proteins, lipid-transfer proteins, hydrogen-peroxide-generating enzymes (Terras *et al.,* 1993; Garcia-Olmedo *et al.,* 1998), stilbene synthase (Hain *et al.,* 1990), ribosome inactivating proteins (Stripe *et al.,* 1992; Longemann *et al.,* 1992), lysozyme from humans, osmotin (PR-5) and osmotin-like protein (Liu *et al.,* 1994; Zhu *et al.,* 1996), polygalacturonase-inhibiting protein, thaumatin and several others. Several herbaceous plants have been engineered with some success by using single genes (*chitinase*, *defensin*, *osmotin*, etc.) or multiple genes (*osmotin* + *chitinase* + *PR1*) (Veronese *et al.,* 1999).

Correlation between the level of expression of antifungal proteins in the leaves and resistance has been observed in several herbaceous transgenes. In field trials the olive plants expressing the *osmotin* gene of tobacco showed reduction of growth (Rugini *et al.,* 2000a; D'Angeli *et al.,* 2001) similarly to the apple plants engineered with the endochitinase gene (Bolar *et al.,* 2000). Research aims at isolating pure compounds (toxins) from fungi, i.e., specific pectic enzymes, malseccin, fusicoccin, fusaric acids, and others to be used as selective pressure on plant cell or tissue culture to recover resistant genotypes, although the resistance acquired by the cells is not always maintained by the derived regenerated plant. However, Orlando *et al.* (1997) demonstrated that pectic enzymes of *Rhizoctonia fragariae* were

effective in selecting strawberry plants resistant to some fungi. Unfortunately the fruits appeared pale red in colour and a little sour, since horto-di-hydroxiphenols in the leaves were increased by over 40%. In the absence of pure compounds the crude culture filtrate of the pathogen could be applied (Hammerschlag and Ognianov, 1990). *In vitro* mutation by ionising rays combined with toxin-used selection seems a promising strategy for future work for fruit crops also.

BACTERIAL RESISTANCE

Every year bacterial diseases cause loss of yield on both tropical and temperate fruit trees. They have effects varying from death of the entire plant to loss of quality of fruits. Important bacterial diseases of fruit trees are fire blight (apple, pear, quince and other ornamental species of *Rosaceae* caused by *Erwiniaamylovora*), bacterial blight and canker of stone fruits (by *Pseudomonas syringae)*, blight of persian walnut(by *Xanthomonas campestris* pv. *Juglandis)* and canker of citruses (by *Xanthomonas citri*). Research on resistance to bacterial diseases has focused on genes producing anti-microbial proteins like lytic peptids (cercopins, attacins and synthetic analogs: shiva-1, SB-37), and lysozymes (egg white, T4 bacteriophage and human lysozyme). Transformation of apple Malling 26 by attacin E (Norelli *et al.,* 1994; Borejsza-Wysocka *et al.,* 1999), and pear cv. Passe Crassane by attacin E and SB-37 (Reynoird *et al.,* 1999a, b; Mourgues *et al.,* 1998) for resistance to *E. amylovora*, are examples of this approach. Recently, relationships betweenattacin expressed in transgenic apple and disease resistance were detected using immunoblot assays with the fusion attacin polyclonal antibody (Ko *et al.,* 1999).

Recent advances in our understanding of *harpin* gene clusters of *P. syringae* and *E. amylovora*, the apoplast conditions for the expression of these genes, their products and secretion systems, and their effects

on host plants, have contributed to clarify the interaction between bacteria and host cells. Other strategies against bacteria effective also in fruit crops are represented by the induction of overproduction of H_2O_2 in the plant cells.

Hydrogen peroxide triggers local hypersensitive cell death, exerts direct antimicrobial activity (Peng and Ku'c, 1992) and is involved in the reinforcement of plant cell wall (Bolwell *et al.,* 1995). Glucose oxidase (*Gox*) gene from *Aspergillus niger*, which induces the production of H_2O_2, increased the level of resistance to *Erwinia carotovora* and *Phytophthora infestans* in potato (Wu *et al.,* 1995) and to *Pseudomonas syringae* and *Xanthomonas campestris* in tomato (Caccia *et al.,* 1999). The resistance orsusceptibility to pathogens can be modified by over expressing hormone genes (Fladung and Gieffers, 1993; Storti *et al.,* 1994). These authors found an increase of resistance to fungi in tomato transgenic plants overexpressing auxin- and cytokinin-synthesising genes (*iaaH* or *iaaM, ipt*) from *Agrobacterium tumefaciens*, when the equilibrium of phytormone of transgenic plants was modified in favour of auxins. Transgenic kiwifruit plants and their transgenic offspring (resulting from crossing *rolABC* staminate cv GTH X normal pistillate Hayward) artificially infected with *Pseudomonas siryngae* and *P. viridiflava*, became more sensitive to these bacteria than untransformed plants (both cv. 'GTH' and 'T1 offspring' noncarrying *rol* genes) (Rugini *et al.,* 1999; Balestra *et al.,* 2001).

NEMATODE RESISTANCE

Many fruit crops are attacked by nematodes of the species *Meloidogyne* spp., *Xiphinema* spp. and *Longidorus* spp. (Brown *et al.,* 1993; Ploetz *et al.,* 1994; Nyczepir and Halbrendt, 1993). Nematodes aredifficult to eradicate from infected soils and control is normally via nematocides, resistant cultivars and appropriate crop husbandry

techniques. However, resistant rootstocks are very rare (Roberts, 1992) and chemical treatments are expensive and not always effective since egg-containing cysts formed by some nematodes are very resistant to chemicals and can survive for years in the soil. Plants respond to infection with a variety of defence strategies including production of phytoalexins, deposition of lignin-like material and accumulation of hydroxyproline- rich glycoproteins, expression of PR-proteins and with an increase of lytic enzymes. Genes involved in nematode resistance have been identified in *Beta procumbens* and *Solanumtuberosum* (*Hs1pro1*and GPA2) and have been cloned (Stiekema *et al.,* 1999).

Two strategies of genetic engineering for introducing resistance to nematodes have been suggested (Sijmons *et al.,* 1994):

1. Introduction of an effector gene whose product is addressed to the parasite or its excretion,
2. Introduction of an effector gene whose product is addressed to the plant cells which feed the nematodes.

Potential anti-nematode genes have been reported and seem to be effective when they are constitutively expressed in plants. Usually these also are involved in the control of insects:

1. Genes over-expressing collagenases which damage the animal cuticle (Havstad *et al.,* 1991).
2. Exotoxin of *B.thuringiensis* (Devidas and Rehberger, 1992) or other feeding inhibitor such as the cowpea trypsin inhibitor.This approach is based on the much localised expression of a phytotoxin gene responsible for the inhibition of development or maintenance of feeding structures of nematodes within the plant system. Genes encoding lipases, transcription factors, nucleases, proteases, and glucanases have been suggested (Sijmons *et al.,* 1994).
3. Anti-nematode monoclonal antibodies (Schots *et al.,* 1992). Molecular information on nematode resistance is limited, the

availability of specific nematode-responsive regulatory plant sequences could represent an important goal and the durability of resistance is believed to depend on the combination of different chimeric constructs and strategies (Barthels *et al.,* 1999).

CONCLUSION

Even when detailed advances in the understanding of the molecular genetics of plants and their pathogens. Whether the use of genetically modified plants will gain public acceptance worldwide is debatable including India. It is also not certain that these approaches will significantly alter the continued arms race between plants and pathogens. The additional selection pressure exerted on pathogens by more elaborate control methods may eradicate some of them, but it may also result in the development of super-pathogens for which yet more elaborate control methods are required. Equally, non-intervention policies have their problems. Apart from reduced yields, many fungi produce mycotoxins, and if these are not adequately controlled, the produce has the potential to be more harmful to human health than if it contains fungicide residues.

Table 1. BACTERIAL RESISTANT TRANSGENIC CROPS

S. No.	Country	Plant name	Trait	Character Gene	Promoter and Terminator	Tolerance to organism	References
1.	China	*Oryza sativa*	Bacterial blight	Xa21	Xa21 Promoter and terminator	*Xanthomonas oryzae* pv. *oryzae* (Xoo),	Lifen Gao *et al.,* 2013
2.	Japan	*B. rapa*	Bacterial wilt	RR S1 and RPS4, *A. thaliana*	*Arabidopsis RRS1* and *RPS4* under the control of their native promoters	*Ralstonia solanacearum*	Mari Narusaka *et al.,* 2014
3.	China	*Oryza sativa*	Bacterial blight	*TaCPK2-A* Calcium-dependent protein kinases (CPKs)	promoters of the A subgenome homologue (*TaCPK2-A*) and D subgenome homoeologue	*Xanthomonas oryzae* pv. *oryzae*, *Xoo*	Shuaifeng Geng *et al.,* 2013

					(*TaCPK2-D*),		
4.	Pakistan	*Oryza sativa*	Bacterial Blight	Xa21	35S promoter	*Xanthomonas oryzae* pv. Oryzae,	Amber Afroz *et al.,* 2012
5.	Pakistan	*Lycopersicon esculentum*	bacterial wilt	Xa21 gene	CaMV 35S promoter	*Pseudomonas solanacearum*	Amber Afroz *et al.,* 2011
6.	Uganda	Banana	Xanthomonas wilt disease	Pflp gene	CaMV35S promoter	*X. campestris* pv. musacearum	Namukwaya *et al.,* 2011
7.	USA	Elite *Indica* rice	Bacterial leaf blight	*Xa21* gene	CaMV 35S promoter	*Xanthomonas oryzae* pv. oryzae	Shiping Zhang *et al.,* 1998
8.	Japan	Rice	Bacterial leaf blight	cecropin B,	CaMV 35S promoter	*Xanthomonas oryzae* pv. oryzae	Arun Sharma *et al.,* 2000
9.	China	Rice	Bacterial leaf blight	Xa26	--	*Xanthomonas oryzae* pv. oryzae	Xinli *et al.,* 2004
10.	Brazil	*Citrus*	Asian citrus	attacin A	CaMV 35S	*Xanthomonas*	Suane

		sinensis L. Osbeck	canker	Gene	promoter	*citri* subsp. citri	Coutinho Cardoso *et al.,* 2010
11.	Bulgaria	Tobacco	Angular leaf spot disease without chlorisis	*ttr* (tabtoxin resistance) gene	CaMV 35S promoter	*Pseudomonas syringae* pv. *tabaci*.	Batchvarova *et al.,* 1998
12.	Japan	Tobacco	Bacterial Diseases	Sarcotoxin IA	CaMV 35S promoter	*Xanthomonas campestris* pv. *Citri, Erwinia carotovora* aubsp. *Carotovora, Pseudomonas syringae* pv. *tabaci*	Masahiro Ohshima *et al.,* 1999

Table 2. FUNGAL RESISTANT TRANSGENIC CROPS

S. No.	Country	Plant name	Trait	Character Gene	Promoter and Terminator	Tolerance to organism	Reference
1	Japan	*Chrysanthemum*	Fungal resistance (Gray mold)	Rice chitinase gene	CaMV 35S promoter EN4	gray mold (*Botrytis cinerea*)	Takatsu *et al.,* 1999
2	USA	Apple	Fungal resistance	Endochitinase from *Trichoderma harzianum*	CaMV 35S promoter	Resistance to Apple Scab and Reduces Vigor	Jyothi Prakash Bolar *et al.,* 2000
3	Australia	Barley	Fungal resistance	Fungal Xylanase gene	*GluB-1* promoter	Assay conformation	Minesh Patel *et al.,* 2000
4	India	*Brassica juncea*	*Alternaria* leaf spot	Class II Chitinase	CaMV35S promoter	*Alternaria brassicae*	Sudesh Chhikara *et al.,* 2012
5	Brazil	Tobacco	Fungal	Chitinase Gene		*Rhizoctonia solani*	Marcelo

			resistance	from *Metarhizium anisopliae*	CaMV 35S promoter		Fernando Kern *et al.,* 2010
6	Japan	Cucumber	Resistance to gray mold	Rice Chitinase gene	CaMV 35S promoter	*Botrytis cinerea*	Tabei et al., 1998
7	Japan	Rice	Blast fungus	Wasabi Defensin gene	Maize ubiquitin-1 promoter	*Magnaporthe grisea*	Kanzaki *et al.,* 2002
8	USA	Cotton	Fungal resistance	Endochitinase gene from *Trichoderma virens*	CaMV 35S promoter	*Alternaria alternata*	Chandrakanth Emani *et al.,* 2003
9	Italy	Tobacco	Fungal resistance	Maize ribosome-inactivating protein b-32	Wun1 promoter, terminator region from the b32.66 cDNA clone	*Rhizoctonia solani*	Massimo Maddaloni *et al.,* 1997

10	Japan	Grapevine	Powdery mildew	Rice Chitinase	CaMV 35S promoter	*Uncinula necator and Elisinoe ampelina*	Yamamoto *et al.,* 2000
11	India	Groundnut	Fungal resistance	Tobacco *β* 1–3 glucanase	35SCaMV promoter and nos terminator	*Cercosporaarachidicola* and *Aspergillusflavus*	Sundaresha *et al.,* 2010
12	Brazil	Lettuce (*Lactuca sativa*)	Fungal resistance	Oxalate decarboxylase gene from *Flammulina* sp.	35S promoter	*Sclerotinia sclerotiorum*	Dias *et al.,* 2006
13	Korea	Rice	Resistance to sheath blight	Maize ribosome-inactivating protein and a rice basic chitinase gene	35S promoter	*Rhizoctonia solani*	Ju-Kon Kim *et al.,* 2003
14	China	Maize	Fungal resistance	Fungal Phytase gene	Maize histone H2B (H2B) promoter, potato	phytase	Rumei Chen *et al.,* 2008

					protease II (PINII) terminator		
15	India	Finger millet (*Eleusine coracana* (L.) Gaertn.)	Leaf blast disease	PIN gene	CaMV 35S promoter	*Pyricularia grisea*	Madhavi Latha *et al.,* 2005
16	Pakistan	*Arachis hypogaea* L.	Resistance Against Leaf Spot	Rice Chitinase Gene	CaMV 35S promoter	Fungal bio assay	Muhammad Munir Iqbal *et al.,* 2012
17	Germany	pea (*Pisum sativum* L.)	Fungal resistance	Chitinase and Glucanase	CaMV 35S promoter	Fungal bio assay	Awah Anna Amian *et al.,* 2011
18	China	*Populus tomentosa* Carr.	Fungal resistance	Chitinase gene (Bbchit1) from Beauveria bassiana	CaMV 35S promoter	*Cytosporachrysosperma*	Zhichun Jia *et al.,* 2010
19	China	*Populus tomentosa* Carr.	Fungal resistance	nsLTP-like antimicrobial protein gene from motherwort	35S CaMV promoter	A. *alternata* (Fr.) Keissler and C. *gloeosporioides* (Penz.),	Zhichun Jia *et al.,* 2010

				(*Leonurusjaponicus*)			
20	China	Potato	Fungal resistance	StoVe1	35S CaMV promoter	Verticillium dahliae	Shui-ping Liu *et al.,* 2012
21	Spain	Rice	Rice blast fungus	Antifungal AFP protein from *Aspergillusgiganteus*	constitutive ubiquitin (ubi) promoter	*Magnaporthegrisea*	Marı´a Coca *et al.,* 2004
22	Japan	Rice	Resistance to the rice blast fungus	Cho1gene	CaMV 35 S promoter	*Magnaportheoryzae*	Yusuke Kouzai *et al.,* 2012
23	Philippines	Rice	Resistance to sheath blight disease	Rice Thaumatin-like protein (PR-5) gene	CaMV 35S promoter	*Rhizoctoniasolani*	Datta *et al .,* 1999
24	India	Tobacco and Peanut	Resistance to	Mustard Defensin gene	CaMV 35S promoter	*Fusariummoniliforme* and	Swathi Anuradha *et*

			fungal pathogens			*Phytophthoraparasitica* pv. *Nicotianae*, *Pheaoisariopsispersonata* and *Cercosporaarachidicola*,	*al.,* 2008
25	Spain	Rice	Resistance to the rice blast fungus	Cecropin A gene	Maize ubiquitin 1 (ubi) promoter, nopaline synthase (nos) terminator sequences	*Magnaporthegrisea*	Marı´a Coca *et al.,* 2006

Table 3. VIRUS, INSECT AND HERBICIDE RESISTANT TRANSGENIC CROPS

S. No.	Country	Plant name	Trait	Character Gene	Promoter and Terminator	Tolerance to organism	Reference
1	Korea	Rice	Resistance against Brown Planthopper	BrD1	Rice cytochrom C promoter and the máåff terminator	*Nilaparvatalugens*	Man-Soo Choi *et al.,* 2009
2	India	Chickpea (*Cicer arietinum* L.)	Insect-resistant	Bt-cry1Ac gene	CaMV35S promoter	*H. armigera*	Meenakshi Mehrotra *et al.,* 2011
3	China	Grape (*Vitis vinifera* L.)	Cold-resistant	AtDREB1b	35S promoter; nos3 = terminator of nitric oxide synthase 3	increase in the conductivity of the transgenic plants was found at –6 _C	Wanmei Jin *et al.,* 2009
4	India	*Chrysanthemum morifolium* cv. Kundan	Virus resistant	Coat protein (CP) gene	CaMV 35S promoter and NOS terminator	Cucumber mosaic virus	Kumar *et al.,*2012
5	USA	Gladiolus	Virus resistant	Coat protein	Arabidopsis	Cucumber mosaic	Kathryn

				subgroup II gene	UBQ3 promoter, NosT	virus	Kamo *et al.,* 2010
6	Japan	Soybean	Herbicide resistant	PAT gene	*Arabidopsis* UBQ3 promoter and NosT	--	Yoichi kita *et al.,* 2009
7	Taiwan	Eustoma grandiflorum	Herbicide resistant	bar gene	CaMV 35S promoter	Basta	Yu-Ting Chen *et al.,* 2010
8	Pakistan	Gossypium hirsutum L.	Virus resistance	tAC1 gene	CaMV35S, CaMV terminator	whitefly-mediated infection	Jamil A. Hashmi *et al.,* 2011
9	Spain	Mexican lime (*Citrus aurantifolia* (Christ.) Swing.)	Virus resistance	*p25* coat protein gene	CaMV 35S promoter and NOS terminator	*Citrus tristeza virus*	Antonio Domínguez, *et al.,* 2002
10	Hungary	*Nicotiana benthamiana*	Virus resistance	Plum pox Virus helicase gene	CaMV 35S promoter and NOS terminator	--	Anita Wittner *et al.,* 1998

11	Greece	Nicotiana benthamiana and Sugar Beet	Beet necrotic yellow vein virus resistance	hrpZ Gene	CaMV35S promoter.	Rhizomania Disease	Ourania *et al.,* 2011
12	Taiwan	Oriental melon	Virus resistance	Coat protein gene	CaMV 35S promoter and NOS terminator	Zucchini yellow mosaic virus	Hui-Wen Wu *et al.,* 2009
13	*Argentina*	Tobacco	Virus resistance	Potato Virus X ORF2 Protein	CaMV 35S promoter and NOS terminator	Tobacco Mosaic Virus and Ob Tobamoviruses	ARES *et al.,* 1998
14	India	Rice	Virus resistance	CP gene	Maize ubiquitin promoter	Rice tungro bacilliform virus	Uma Ganesan *et al.,* 2009
15	Canada	Rice	Pest resistant	cry1Ca1 Gene	35S cauliflower mosaic virus (d35S) promoter, NOS T	*Spodoptera litura* and *Chilo suppressalis*	Mohsin Abbas Zaidi *et al.,* 2009
16	USA	*Oryza sativa* L.	Virus resistance	Rice tungro spherical	Rice actin promoter	--	Elumalai Sivamani *et*

				virus (RTSV) coat protein	and nopaline synthase T		*al.,* 1999
17	Japan	Soybean	Virus resistance	sense coat protein gene	CaMV 35S promoter and NOS terminator	Soybean dwarf virus-resistant	Makoto Tougou *et al.,* 2007
18	Singapore	Sugarcane	Insect resistant	cry1Ac	Maize Ubi-1 promoter	Control against stem borers	Li-Xing Weng *et al.,* 2011
19	Cuba	Sugarcane	Insect resistant	*Bacillus thuringiensis* _-endotoxin	CaMV 35S promoter, Tnos terminator.	Resistant to stem borer attack	Ariel Arencibia *et al.,* 1997
20	Italy	Tomato	Virus resistance	Truncated Viral Rep Protein	35S promoter	Resistance to Tomato Yellow Leaf Curl Virus	Brunetti *et al.,* 1997
21	Taiwan	Watermelon	Virus resistance	Partial CP genes	Cauliflower mosaic virus (CaMV) 35S promoter	Resistant to Cucumber mosaic virus and Watermelon mosaic virus	Ching-Yi Lin *et al.,* 2011

References

1. Abdul Kader, A.M.; Aldwinckle, H.S.; Bauer, D.W.; Beer, S.V. and Norelli, J.L. (1999). 'Evaluation of the *hrp N* gene for increasing resistance to fire blight in transgenic apple'. *Acta Hort.,* **489**: 247-250.
2. Agrawal, N.; Dasaradhi, P.V.; Mohmmed, A., Malhotra, P.; Bhatnagar, R.K. and Mukherjee, S.K. (2003). RNA interference: biology, mechanism, and applications. *Microbiology and Molecular BiologyReviews*, **67**: 657–685.
3. Alrefai, R.H. and Korban, S.S. (1995). Cross protection against virus diseases in fruit trees. *Fruit. Var. J.,***49**: 21–30.
4. Aziz, R. and Tepfer, M. (1999). 'Recombination in RNA viruses and in virus-resistant transgenic plants'. *J.General Virology,* **8**: 1339–1346.
5. Balestra, G.M.; Rugini, E. and Varvaro, L., (2001). 'Increasing of susceptibility to *Pseudomonas syringae* pv. *syringae* and *Pseudomonas viridiflava* of transgenic *rolABC* genes kiwi plants and its inheritance in T1 offspring'. *J. Phytopathol.,***149**: 189–194.
6. Barbieri, L.; Valbonesi, P.; Bonora, E.; Gorini, P.; Bolognesi, A. and Stirpe, F. (1997). 'Polynucleotide: adenosine glycosidase activity of ribosome inactivating proteins: Effect on DNA, RNA, and poly-(A)'. *Nucl.Ac. Res*.,**25**: 518–522.
7. Barcelo, P.; Hagel, C.; Becker, D.; Martin, A. and Lo Rz, H. (1994). Transgenic cereal (tritordeum) plants obtained at high efficiency by microprojectile bombardment of inflorescence tissue. *Plant J.,***5:** 583–592.
8. Barthels, N.; Karimi, M.; Vercauteren, I.; Van Montagu, M. and Gheysen, G. (1999). Integration of nematode-responsive regulatory sequences from *Arabidopsis thaliana* into nematode control strategies. In: G.T. Scarascia Mugnozza, E. Porceddu, and

M.A. Pagnotta ,(eds.) *Genetics and Breeding forCrop Quality and Resistance.* Kluwer Academic Publishers, The Netherlands, pp. 203–210.

9. Baulcombe, D. (1994). 'Replicase-mediated resistance. A novel type of transgenic resistance in transgenic plants'. *Trends in Microbiology.,* **2**: 60–63.
10. Bevan, M.; Flavell, R.B. and Chilton, M.D. (1983). A chimeric antibiotic resistance gene as selectable marker for plant cell transformation. *Nature,* **394:** 184–187.
11. Bolar, J.P.; Norelli, J.L.; Wong, K.W.; Hayes, C.K.; Harman, G.E. and Aldwinckle, H.S. (2000). 'Expression of endochitinase from *Tricoderma harzianum* in transgenic apple increases resistance to apple scab and reduces vigour'. *Phytopathology*, **90**: 72–77.
12. Bolwell, G.P.; Butt, V.S.; Davies, D.R. and Zimmerlin, A. (1995). 'The origin of the oxidative burst in plants'.*Free Rad. Res. Comm.,* **23**: 517–532.
13. Borejsza-Wysocka, E.E.; Norelli, J.L. and Ko, K. (1999). 'Transformation of authentic M. 26 apple rootstock for enhanced resistance to fire blight'. *Acta Hort.,***489**: 259–266.
14. Bourdin, D. and Lecoq, H. (1991). 'Evidence that heteroencapsidation between two polyviruses is involved in aphid transmision of a non-aphid-transmissibile isolate from mixed infections'. *Phytopathology*, **81**: 1459–1464.
15. Broekaert, W.F.; Terras, F.R.G.; Cammue, B.P.A. and Osborn, R.W. (1995) . 'Plant defensins: novel antimicrobial peptides as components of the host defense system.' *Plant Physiol.,***108**: 1353–1358.
16. Brown, D.J.F.; Dalmasso, A. and Trudgill, D.L. (1993). 'Nematode pests of soft fruits and vines'. In: K. Evans, D.L. Trudgill, and J.M. Webster (eds). *Plant parasitic nematodes in temperate*

agriculture. CAB Int., Wallingford, UK, pp. 427–462.

17. Buzkan, N.; Minafra, K.; Saidarelli, P.; Castellano, M.K. and Martelli, G.P. (2000). 'Heteroencapsidation in transgenic and non-transgenic *Nicotiana* plants infected by Grapevine virus A and B'. *ExtendedAbstracts*, 13th Meeting 1CVG, Adelaide 2000.
18. Caboche, M. (1990). Liposome-mediated transfer of nucleic acids into plant cells. *Physiol Plant.,***79:** 173– 176.
19. Caccia, R.; Delledonne, M.; Balestra, G.M.; Varvaro, L. and Soressi, G.P. (1999). Plant-bacterial pathogen interaction modified in transgenic tomato plants expressing the Gox gene encoding glucose oxidase. In: G.T. Scarascia Mugnozza, E. Porceddu and M.A. Pagnotta (eds) *Genetics and Breeding forCrop Quality and Resistance*. Kluwer Academic Publishers, The Netherlands, pp., 119–125.
20. Carr, P. and Zaitlin, M. (1993). 'Replicase-mediate resistance'. *Seminars in Virology,***4**: 330–347.
21. Chowrira, G.M.; Akella, V. and Lurquin, P.F. (1995). Electroporation-mediated gene transfer into intact nodal meristems in Planta. *Molecular Biotechnology,***3:** 17–23.
22. Christou, P. (1992). Genetic transformation of crop plants using microprojectile bombardment. *Plant J.,***2:** 275–281.
23. Christou, P.; Ford, T. and Kofron, M. (1991). Production of transgenic rice (Oriza sative L.) plants from agronomically important indica and japonica varieties via electric discharge particle acceleration of exogenous DNA into immature zygotic embryos. *Bio Tech.,***9:** 957–962.
24. Christou, P.; Murphy, J.E. and Swain, W.F. (1987). Stable transformation of soybean by electroporation and root formation from transformed callus. *Proc. Natl. Acad. Sci.,***84:** 3962–3966.

25. Cogoni, C. and Macino, G. (2000). Gene silencing across kingdoms. *Current Opinion in GeneticsDevelopment,* **10**: 638–643.
26. Cornelissen, B.J.C. and Melchers, L.S. (1993). Strategies for control of fungal diseases with transgenic plants. *Plant Physiol.,* **101**: 709–712.
27. Creamer, R. and Falk B.W. (1990). Direct detection of transcapsidated barley yellow dwarf luteovirus in double infected plants. *J. Gen. Virology,***71**: 211–217.
28. D'angeli, S.; Gutierrez-Pesce, P.; Altamura, M.M.; Biasi, R.; Ruggiero, B.; Muganu, M.; Bressan, R. and Rugini, E. (2001). 'Genetic transformation of olive tree (*Olea europaea* L.) with *osmotin* gene and *in situ* protein localisation in the transgenic tissues'. Italian Society of Agriculture Genetics, XLVAnnual Congress, Salsomaggiore Terme (PR) Italy.
29. D'halluin, K.; Bonne, E.; Bossut, M.; De Beuckeleer, M. and Leemans, J. (1992). Transgenic maize plants by tissue electroporation. *Plant Cell,***4:** 1495–1505.
30. De Block, M.; Herrera-Estrella, L.; Van Montagu, M.; Schell, J. and Zambryski, P. (1984). Expression of foreign genes in regenerated plants and their progeny. *EMBO J.,***3:** 1681–1689.
31. De Wet, J.; Wood, K.V.; Helinski, D.R. and Deluca, M. (1985). Cloning of firefly luciferas cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci., USA,***82:** 7870–7873.
32. Delledonne, M.; Xia Y.; Dixon, R.A.; Lorenzoni, C. and Lamb, C. (1999). Nitric oxide signalling in the plant hypersensitive disease resistance response. In: G.T. Scarascia Mugnozza, E. Porceddu and M.A. Pagnotta (eds) *Genetics and Breeding for Crop Quality and Resistance.* Kluwer Academic Publishers, The Netherlands,

pp. 127–133.

33. Devidas, P. and Rehberger, L.A. (1992). 'The effect of exotoxin from *Bacillus thuringiensis* on *Meloidogyneincognita* and *Caenorhabditis elegans*'. *Plant & Soil,* **145**: 115–120.

34. Dolgov, S.V.; Lebedev, V.G.; Anisimova, S.S.; Lavrova, N.; Serdobinskiy, L.A.; Tjukavin, G.B.; Shadenkov, S.A. and Lunin, V.G. (1999). Phytopathogen resistance improvement of horticultural crops by plant-defensin gene introduction'. In: G.T. Scarascia Mugnozza, E. Porceddu and M.A. Pagnotta (eds) *Genetics and Breeding for Crop Quality and Resistance.* Kluwer Academic Publishers, TheNetherlands, pp., 111–118.

35. Dolja V.V.; Halderman R.; Robertson N.L.; Dougherty W.G. and Carrington J.C. (1994). 'Distinct functions of capsid protein in assembly and movement of tobacco etch potyvirus in plants'. *EMBO J.,***13**: 1482–1491.

36. Dominguez, A.; Guerri, J.; Cambra, M.; Navarro, L.; Moreno, P. and Pegna, L. (2000). 'Efficient production of transgenic citrus plants expressing the coat protein gene of citrus tristeza virus'. *Plant Cell Rep.,***19**: 427–433.

37. Dong, Y. and Friedrich, M. (2005). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**: 806–811.

38. Farinelli, L.; Malnoe, P. and Collet, G.F. (1992). Heterologous encapsidation of potato virus Y strain 0 (PVYO) with the transgenic coat protein of PVY strain N (PVYN) in *Solanum tuberosum* cv. Bintje. *Biotechnol.* **10**: 1020–1025.

39. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E. and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature, **391**: 806-811.

40. Fitchen, J.H. and Beachy, R.N. (1993). Genetically engineered protection against viruses in transgenic plants. *Annu. Rev. Microbiol.,***47**: 739–763.

41. Fladung, M. and Gieffers, W. (1993). Resistance reactions of leaves and tubers of *rol* C transgenic tetraploid potato to bacterial and fungal pathogens. Correlation with sugar, starch and chlorophyll content. *Phys. and Mol. Plant Path*.,**42**: 123–132.

42. Fraley, R.T.; Rogers, S.G.; Horsch, R.B.; Sanders, P.R.; Flick, J.S.; Adams, S.P.; Bittner, M.L.; Brand, L.A.; Fink, C.L.; Fry, J.S.; Galluppi, G.R.; Goldberg, S.B.; Hoffmann, N.L. and Woo, S.C. (1983). Expression of bacterial genes in plant cells. *Proc. Natl. Acad. Sci., USA,***80:** 4803–4807.

43. Fraley, R.T.; Rogers, S.G. and Horsch, R.B. (1986). *Crit. Rev. Plant. Sci.,***4:** 1–45.

44. Fromm, M.E.; Taylor, L.P. and Walbot, V. (1986). Stable transformation of maize after electroporation. *Nature,* **319:**791–793.

45. Fromm, M.; Callis, J.; Taylor, L.P. and Walbot, V. (1987). Electroporation of DNA and RNA into plant protoplasts. *Methods Enzymol,* **153:** 351–366.

46. Gadani, F.; Mansky, L.M.; Medici, R.; Mille, W.A. and Hull, J.N. (1990). Genetic engineering of plants for virus resistance. *Arch. Viral.,* **115:** 1–21.

47. Garcia-Olmedo, F.; Molina, A.; Alamillo, J.M.; Rodriguez Palenzuela, P. (1998). Plant defense peptides. *Biopolymers,* **47**: 479–491.

48. Gonsalves, D. (1998). Control of papaya ringspot virus in papaya: a case study'. *Annual Review ofPhytopathology,* 415–437.

49. Greene A.E. and Allison R.F. (1994). Recombination between viral RNA and transgenic plant transcripts. *Science,* **263**: 1423–

1425.

50. Grumet, R. (1994). Development of virus resistance plants via genetic engineering. *Plant Breed. Rev.,***12:** 47–79.
51. Guerche, P.; Bellini, C.; Lemoullec, J.M. and Cabache, M. (1987). *Biochimie,***69:** 621–628.
52. Hahn, M.G.; Bucheli, P.; Cervone, F.; Doares, S.H.; O'neill, R.A.; Darvill, A. and Albersheim, P. (1989). Roles of cell wall constituents in plant-pathogen interactions. *Plant-Microbe Interactions,***3:** 131– 181.
53. Hain, R.; Bieseler, B.; Kindel, H.; Schroder, G. and Stocker, R. (1990). Expression of a stilbene sythase gene in *Nicotiana tabacum* results in synthesis of the phytoalexin resveratron. *Plant Mol. Biol.,***15**: 325–335.
54. Haldrup, A.; Petersen, S.G. and Okkels, F.T. (1998). The xylose isomerase gene from *Thermoanaerobacterium thermosulforogenes* allows effective selection of transgenic plant cellsusing D-xylose as the selection agent. *Plant Mol. Biol.,***37:** 287–296.
55. Hammerschlag, F.A. and Ognianov, V. (1990). Somaclonal variation in peach: Screening for resistance to *Xanthomonas campestris* pv. *pruni* and *Pseudomonas syringae* pv. *syringae*. *Acta Hort.,* **280**:403–408.
56. Hammond-Kosack, K.E. and Jones, J.D.G. (1996). Plant disease resistance genes. *Annu. Rev. Plant Physiol.Plant Mol. Biol.,* **48**: 575–607.
57. Hannon, G.J. (2002). RNA interference. *Nature*, **418**: 244–251.
58. Havstad, P.; Sutton, D.; Thomas, S.; Sengupta-Gopalan, C. and Kemp, J. (1991). Collagenase expression in transgenic plants: an alternative to nematicides. *Mol. Biol. Plant Growth Develop.,* p. 354 (abstract).

59. Hayashimoto, A.; Li, Z. and Murai, N. (1990). A PEG-mediated protoplast transformation system for production of fertile transgenic rice plants. *Plant Physiol.,***93:** 857–863.

60. Heim, R.; Cubitt, A.B. and Tsien, R.Y. (1995). Improved green fluorescence. *Nature,***373:** 663–664.

61. Hille, J.; Verheggen, F.; Roelvink, P.; Franssen, H.; Kammen, A.V. and Zabel, P. (1986). *Plant Mol Biol.,***7:** 171–176.

62. Hobbs H.A. and Mclaughlin M.R. (1990). A non aphid-transmissible isolate of bean yellow mosaic virus-Scott that is transmissibile from mixed infections with pea mosaic virus-204-1. *Phytopathology,***80**: 268–272.

63. Jacquemond, M. and Tepfer, M. (1998). Satellite RNA-mediated resistance to plant viruses: are the ecological risks well assessed. In: A. Hadidi, R.H. Khetarpal and H. Koganezawa (eds) *Plant Virus DiseaseControl*. American Phytopathological Society Press, St. Paul, pp., 94–120.

64. Jefferson, R.A.; Burgess, S.M. and David, H. (1986). ß -glucuronidase from *Escherichia coli* as a gene-fusion marker. *Proc. Natl. Acad. Sci., USA,***83:** 8447–8451.

65. Joersbo, M. and Okkels, F.T. (1996). A novel principle for selection of transgenic plant cells: positive selection. *Plant Cell Rep.,* **16:**219–221.

66. Kaether, C. and Gerdes, H.H. (1995). Visualization of protein transport along the secretory pathway using green fluorescence protein. *FEBS Lett.,***369:** 267–271.

67. Kaniewski, W.K. and Lawson E.C. (1998). Coat protein and replicase-mediated resistance to plant viruses. In: A. Hadidi, R.H. Khetarpal and H. Koganezawa (eds) *Plant Virus Disease Control*. American Phytopathological Society Press, St. Paul, pp., 65–78.

68. Keen, N.T. and Yoshikawa, M. (1983). ß-1, 3 endoglucanase from

soybean releases elicitor-active carbohydrates from fungus cell walls. *Plant Physiol.,***71:** 460–465.

69. Khetarpal, R.K. ; Maisonneuve, B. ; Maury, Y. ; Chalhoub, B. ; Dinant, S. ; Lecoq, H. and Varma, A. (1998). Breeding for resistance to plant viruses. In: A. Hadidi, R.H. Khetarpal and H. Koganezawa (eds) *Plant Virus Disease Control*. American Phytopathological Society Press, St. Paul, pp., 1–32.
70. Klein, T.M.; Wolf, E.D.; Wu, R. and Sanford, J.C. (1987). High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature,* **327:** 70–73.
71. KO, K.; Norelli, J.L.; Brown, S.K.; Aldwinckle, H.S. and During, K. (1999). Galaxy line transgenic for attacin E and T4 lysozyme genes has increased resistance to fire blight'. In: A. Altman, M. Ziv and S. Izhar (Eds) *Plant Biotechnology and In Vitro Biology in the 21st Century*, vol. **36**. Kluver Academic Publishers, Dordrecht, pp. 507–511.
72. Lamb, C. and Dixon, R.A. (1997). The oxidative burst in plant disease resistance. *Annu. Rev. Plant Physiol.Plant. Mol. Biol.,* **48**: 251–275.
73. Lamb, C.J.; Ryals, J.A.; Ward, E.R. and Dixon, R.A. (1992). 'Emerging strategies for enhancing crop resistance to microbial pathogens. *Biotechnol*. **11:** 1436–1445.
74. Liu, D.; Raghothama, K.G.; Hasegawa, P.M. and Bressan, R.A. (1994). 'Osmotic over expression in potato delays development of disease symptoms. *Proc. Natl. Acad. Sci., USA,***91:** 1888–1892.
75. Longemann, J.; Jack, G.; Tommerup, H.; Munday, J. and Schell, J. (1992). Expression of barley ribosome-inactivating protein leads to increased fungal protection in transgenic tobacco plants. *Biotechnol.,***10:** 305–308.
76. Maddaloni, M.; Forlani, F.; Balmas, V.; Donini, G.; Corazza, L.;

Fang, H.; Pincus, S. and Motto, M. (1999). The role of b-32 protein in protecting plants against pathogens'. In: G.T. Scarascia Mugnozza, E. Porceddu and M.A. Pagnotta (eds) *Genetics and Breeding for Crop Quality and Resistance.* Kluwer Academic Publishers, The Netherlands, pp. 77–82.

77. Martelli, G.P. ; Gallitelli, D. and Russo, M. (1999). An appraisal of pathogen-derived resistance for the control of virus diseases. In: G.T. Scarascia Mungozza, E. Porceddu, M.A. Pagnotta (eds) *Geneticsand Breeding for Crop Quality and Resistance*. Kluwer Academic Publishers, Dordrecht, pp.,223–232.
78. Mauch, F.; Hadwiger, L.A. and Boller, T. (1988). Antifungal hydrolases in pea tissues. I. Purification and characterisation of two chitinases and two ß- 1, 3-glucanases differentially regulated during development and in response to fungal infection. *Plant Physiol*., **87**: 325–333
79. Melchers, L.S.; Ponstein, A.S.; Sela-Buurlage, M.B.; Vloemans, S.A. and Cornelissen B.J.C. (1993). In vitro anti-microbial activities of defence proteins and biotechnology'. In: B. Fritig and M. Legrand (eds) *Mechanisms of plant defense responses.* Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 401–410.
80. Mikkelsen, J.D.; Berglund, L.; Nielson, K.K.; Christiansen, H. and Bojsen, K. (1992). Structure of endochitinase genes from sugar beets'. In: C.J. Brine, P.A. Sanford, and J.P. Zikakis (eds), *Advancesin chitin and chitosan.* Elsevier Science Publ., London, UK, pp. 344–35.
81. Miller, W.A.; Koev, G. and Mohan, B.R. (1997). Are risks associated with transgenic resistance to luteoviruses? *Plant Disease,* **81:**700–710.
82. Mourgues, F.; Brisset, M.N. and Chevreau, E. (1998). Strategies to improve plant resistance to bacterial diseases through genetic

engineering. *Trends Biotechnology*, **16:** 203–210.

83. Nelson, R.S.; Mccormick, S.M.; Delannay, X.; Dube, P.; Layton, J.; Anderson, E.J.; Kaniewska, M.; Proksch, R.K.; Horsch, R.B.; Rogers, S.G.; Fraley, R.T. and Beachy, R.N. (1988). Virus tolerance, plant growth, and field performance of transgenic tomato plants expressing coat protein from tobacco mosaic virus. *Bio Technology,***6:** 403–409.

84. Niwa, Y.; Hirano, T.; Yoshimoto, K.; Shimizu, M. and Kobayashi, H. (1999). Noninvasive quantitative detecting and applications of non-toxic, S65T-type green fluorescent protein in living plants. *Plant J.,***18** 455– 463.

85. Norelli, J.L.; Aldwinckle, H.S.; Destefano-Beltran, L. and Jaynes, J.M. (1994). 'Transgenic ''Malling 26'' ['M. 7''] apple expressing the *attacin E* gene has increased resistance to *Erwinia amylovora*'. In: H. Schmidt and M. Kellerhals (eds), *Progress in temperate fruit breeding.* Kluwer Academic Publishers, London, UK, pp. 333–338.

86. Nyczepir, A.P. and Halbrendt J.M. (1993). Nematode pests of deciduous fruit and nut trees. In: K. Evans, D.L. Trudgill, and J.M. Webster (eds), *Plant parasitic nematodes in temperate agriculture.* CAB Int., Wallingford, UK, pp. 381–425.

87. Orlando, R.; Magro, P. and Rugini, E. (1997). Pectic enzyme as a selective press, for in vitro recovering of strawberry plants with a multiple fungal disease resistance. *Plant Cell Rep.,***16:** 272–276.

88. Peng, M. and Ku'c, J. (1992). Peroxidase-generated hydrogen peroxide as source of antifungal activity in vitro and on tobacco leaf disks. *Phytopathology,***82:** 696–699.

89. Ploetz, R.C.; Zentmyer, G.A.; Nishijia, W.T.; Rohrbach, K.G. and Ohr, H.D. (1994). *Compendium of tropicalfruit diseases.* APS Press, St. Paul, MN. pp. 1–84.

90. Reynoird, J.P. ; Mourgues, F.; Chevreau, E. and Aldwinckle, H.S. (1999). Expression of SB-37 gene in transgenic pears enhanced resistance to fire blight'. *Acta Hort.,***489:** 243–244.

91. Reynoird, J.P. ; Mourgues, F.; Chevreau, E. and Brisset, M.N. (1999b). First evidence of fire blight resistance among transgenic pear clones expressing *attacin E* gene. *Acta Hort.,***489:** 245–246.

92. Riggs, C.D. and Bates, G.W. (1986). Stable transformation of tobacco by electroporation: evidence for plasmid concatenation. *Proc. Natl. Acad. Sci., USA,***83:** 5602–5606.

93. Roberts, P.A. (1992). Current status of the availability, development and use of host plant resistance to nematodes. *J. Nematol.,***24**: 213–223.

94. Robinson, D.; Taliansky, M. and Ryabov, E. (1999). Expression of luteovirus coat protein in transgenic plants revealed by infection with an umbravirus. *XI Int. Cong. Virology*, Sydney, p. 49.

95. Roossinck, M.J. (1997). Mechanisms of plant virus evolution. *Annual Review of Phytopathology,***35:** 191–209.

96. Rouwendal, G.J.; Mendes, O.; Wolbert, E.J. and Douwe, D.E.B. (1997). Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage. *Plant Mol. Biol.,***33:** 989–999.

97. Rugini, E.; Biasi, R. and Muleo, R. (2000a). Olive (*Olea europaea* var. *sativa*) Transformation'. In: S.M. Jain and S.C. Minocha (eds), *Molecular Biology of Woody Plants, vol., 2* Kluwer Academic Publishers, pp. 245–279.

98. Rugini, E.; Muganu, M.; Pilotti, M.; Balestra, G.M.; Varvaro, L.; Magro, P.; Bressan, R. and Taratufolo, C. (1999). Genetic stability, transgene hereditability and agronomic evaluation of transgenic kiwi (*Actinidia deliciosa* A. Chev.) plants for *rolABC*, *rolB* and Osmotin genes. *Fourth InternationalSymposium on*

Kiwifruit, Santiago del Chile 11–14 Febr., p. 26.

99. Rybicki, E.P. (1994). A phylogenetic and evolutionary justification for the three genera of Geminiviridae. *Archives of Virology,* **139:**49–77.

100. Sanford, J.C. and Johnston, S.A. (1985). The concept of parasite-derived resistancederiving resistance genes from the parasite's own genome. *J. Theor. Biol.,***113:** 395–405.

101. Schnorf, M.; Neuhaus -Uri, G.; Galli, A.; Iida, S. and Potrykus, I. (1991). An improved approach for transformation of plant cells by microinjection: molecular and genetic analysis. *Transgenic Res.,***1:** 23–30.

102. Schots, A.; Deboer, J.; Schouten, A.; Roosein, J. and Zilverentant, J.F. (1992). Plantibodies: a flexible approach to design resistance against pathogens. *Neth. J. Plant Pathol.,***98:** 183–191.

103. Shimamoto, K.; Terada, R.; Izawa, T. and Fujimoto, H. (1989). Fertile transgenic rice plants regenerated from transformed protoplasts. *Science,***338:** 274–276.

104. Sijmons, P.C.; Atkinson, H.J.; and Wyss, U. (1994). 'Parasitic strategies of root nematodes and associated host all responses. *Annu. Rev. Phytopathol.,***32:** 235–259.

105. Singh, Z.; Jones, R.A.C. and Jones, M.G.K. (1997). Effectiveness of coat protein and defective replicase gene mediated resistance against Australian isolates of cucumber mosaic cucumovirus'. *Proc. 11thBiennial Conference Australasian Plant Path. So.,* p. 271.

106. Smith, G.R.; Borg, Z.; Lockhart, B.E.L.; Braithwaite, C.S. and Gibbs, M.J. (2000). Sugarcane yellow leaf virus: a novel member of the Luteoviridae that probably arose by inter-specific recombination. *J.General Virology,* **81:**1865–1869.

107. Smyth, D.R. (1999). Gene silencing: plants and viruses fight it out.

Current Biology, **9:** 100–102.

108. Soreq, H. and Seidman, S. (2001). Acetylcholinesterase – new roles for an old actor. *Nature ReviewNeurosciences*,**2**: 8–17.

109. Staskawicz, B.J.; Ausubel, F.M.; Baker, B.J.; Ellis, J.G. and Jones, J.D.G. (1995). Molecular genetics of plant disease resistance. *Science,* **268:** 661–667.

110. Stiekema, V.J.; Van Der Vossen, A.G.; Rouppe Van Der Voot, J.; Bakker, J. and Klein Lankhorst, R.M. (1999). Molecular isolation of two cyst nematode resistance genes: the Hs1pro-1 gene of beet and the GPA2 gene of potato. In: G.T. Scarascia Mugnozza, E. Porceddu and M.A. Pagnotta (eds) *Genetics and Breeding for Crop Quality and Resistance.* Kluwer Academic Publishers, TheNetherlands, pp. 185–193.

111. Storti, E.; Bogani, P.; Bettini, P.; Bittini, P.; Guardiola, M.L.; Pellegrini, M.G.; Inze, D. and Buiatti, M. (1994). Modification of competence for in vitro response to *Fusarium oxysporum* in tomato cells. II. Effect of the integration of *Agrobacterium tumefaciens* genes for auxin an cytokinins synthesis. *Theor.Appl. Genet.,* **88:**89–96.

112. Stripe, F.; Barbieri, L.; Battelli, L.G.; Soria, M. and Lappi, D.A. (1992). Ribosome inactivating proteins from plants, present status and future prospects. *Biotechnol.,* **10:** 405–412.

113. Takeuchi, Y.; Yoshikawa, M.; Taekba, G.; Tanaka, K.; Shibata, D. and Horino, O. (1990). Molecular cloning and ethylene induction of mRNA encoding a phytoalexin elicitor releasing factor, ß-1,3 endoglucanase, in soybean. *Plant Physiol.*, 673–682.

114. Tennant, P.R.; Consalves, C.; Ling, K.S.; Fitch, M.; Manshardt, R.; Slightom, J.L. and Gonsalves, D. (1994). Differential protection against papaya ringspot virus isolates in coat protein gene transgenic papaya and classically crossprotected papaya.

*Phytopathology,***84:** 1359–1366.

115. Terras, F.R.; Penninckx, I.A.; Goderis, I.J. and Broekaert, W.F. (1998). Evidence that the role of plant defensins in radish defense responses is independent of salicylic acid. *Planta*, **206:** 117–124.
116. Terras, F.R.; Torrekens, S.; Van Leuven, F.; Osborn, R.W.; Vanderleyden, J.; Cammue, B.P. and Broeckaert, W.F. (1993). A new family of basic cysteinerich plant antifungal proteins from Brassicaceae species. *FEBS Lett.,* **316:**233–240.
117. Torres, M.; Siemens, J.; Meixner, M. and Sacrista, N.M.D. (1997). An improved method for direct gene transfer and subsequent regeneration of *Arabidopsis thaliana* Landsberg erecta and two marker lines. *Plant Cell Tissue and Organ Culture,***47:** 111–118.
118. Vancanneyt, G.; Schmidt, R.; O'connor-Sanchez, A.; Willmitzer, L. and Rocha-Sosa, M. (1990). Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in *Agrobacterium*-mediated plant transformation. *Mol. Gen. Genet.,***220:** 245–250.
119. Veronese, P.; Crino, P.; Tucci, M.; Colucci, F.; Yun, D.J.; Hasegawa, M.P.; Bressan, R.A. and Saccardo, F. (1999). 'Pathogenesis-related proteins for the control of fungal diseases of tomato. In: G.T. Scarascia Mugnozza, E. Porceddu, M.A. Pagnotta (eds) *Genetics and Breeding for Crop Quality andResistance.* Kluwer Academic Publishers. Netherlands, pp. 15–24.
120. Waterworth, H.E. and Hadidi, A. (1998). Economic losses due to plant viruses. In: A. Hadidi, R.H. Khetarpal and H. Koganezavra (eds) *Plant Virus Disease Control* 1–13. American Phytopathological Society Press, St. Paul.
121. Wilson, T.M.A. (1993). Strategies to protect crop plants against

viruses'. *Proc. Nat. Acad. Sci., USA,* **90:** 3134–3141.

122. Wu, G.S.; Short, B.J.; Lawrence, E.B.; Levine, E.B.; Fitzsimmons, K.C. and Shah, D.M. (1995). Disease resistance conferred by expression of a gene encoding H_2O_2 -generating glucose oxidase in transgenic potato plants. *Plant Cell.,* **7:** 1357–1368.

123. Yang, Z.N.; Ingelbrecht, I.L.; Louzada, E.; Svaria, M. and Mirkov, T.E.; (2000). *Agrobacterium*-mediated transformation of the commercial important grapefruit cultivar Rio Red (*Citrus paradisi* Macf.). *Plant Cell Rep.,* **19:**1203–1211.

124. Yie, Y. and Tien, P. (1998). Controlling mosaic virus diseases under field conditions using multiple gene strategies in transgenic plants'. In: A. Hadidi, R.H. Khetarpal and H. Koganezawa (eds) *PlantVirus Disease Control.* American Phytopathological Society Press, St. Paul, pp. 129–141.

125. Zhang, L.Y. and Shi, Y.F. (2002). Induction of acetylcholinesterase expression during apoptosis in various cell types. *Cell Death and Differentiation,* **9**: 790–800.

Chapter 6
Animal Biotechnology

Shilpa R Raju, Suma M.S, Nalina M and
Chandrashekara K.N

Animal biotechnology is a broad field encompassing the polarities of fundamental and applied research, including molecular modeling, gene manipulation, development of diagnostics and vaccines and manipulation of tissue. It accounts for the use of biotechnology tools, including molecular markers, stem cells, and tissue engineering. Molecular markers are increasingly being used to identify and select the particular genes that lead to desirable traits and it is now possible to select superior germ plasma and disseminate it using artificial insemination, embryo transfer and other assisted reproductive technologies. These technologies have been used in the genetic improvement of livestock. Transgenesis offers considerable opportunity for advances in medicine and agriculture. In livestock, the ability to insert new genes for such economically important characteristics as fecundity, resistance to or tolerance of other environmental stresses would represent a major breakthrough in the breeding of commercially superior stock. Another opportunity that transgenic technology could provide is in the production of medically important proteins such as insulin and clotting factors in the milk of domestic livestock. A comprehensive evaluation of strategies for developing, testing, breeding and disseminating transgenic livestock in the context of quantitative

improvement of economic traits is being done. Genetic improvement of livestock depends on access to genetic variation and effective methods for exploiting this variation. Genetic diversity constitutes a buffer against changes in the environment and is a key in selection and breeding for adaptability and production in a range of environments. Animal cell culture technology in today's scenario has become indispensable in the field of life sciences, which provides a basis to study regulation, proliferation, differentiation, and to perform genetic manipulation. It requires specific technical skills to carry out successfully. Application of tissue culture includes the study and understanding of intracellular activity, intracellular flux, pharmacology, cell-cell interaction, cell products, toxicology, tissue engineering, genomics, and immunology. Knowledge acquired from these studies can be used in the biomedical applications.

- **Culture Media:** The culture medium is the most important component of the culture environment, because it provides the necessary nutrients, growth factors, and hormones for cell growth, as well as regulating the pH and the osmotic pressure of the culture. Although initial cell culture experiments were performed using natural media obtained from tissue extracts and body fluids, the need for standardization, media quality, and increased demand led to the development of defined media. The three basic classes of media are basal media, reduced-serum media, and serum-free media, which differ in their requirement for supplementation with serum.
- **Media Components Balanced Salt Solutions:** A balanced salt solution (BSS) is composed of inorganic salts and may include sodium carbonate and, in some cases, glucose. Commercial complete media will list which BSS formulation was used.
- **Serum:** Serum is vitally important as a source of growth and

adhesion factors, hormones, lipids and minerals for the culture of cells in basal media. In addition, serum also regulates cell membrane permeability and serves as a carrier for lipids, enzymes, micronutrients, and trace elements into the cell. However, using serum in media has a number of disadvantages including high cost, problems with standardization, specificity, variability, and unwanted effects such as stimulation or inhibition of growth and/or cellular function on certain cell cultures. If the serum is not obtained from reputable source, contamination can also pose a serious threat to successful cell culture experiments. Always check new batches of serum before use. The quality and the composition can vary greatly from batch to batch. Serum is inactivated by incubating it for 30 min at +56^{o}C. Originally, heating was used to inactivate complements for immunoassays, but it may also have other, undocumented effects.

- **Other Supplements:** In addition to serum, tissue extracts and digests have traditionally been used to supplement tissue culture media. The most common ones are amino acid hydrolysates (from beef heart) and embryo extract (chick embryo).
- **Basal Media**: The majority of cell lines grow well in basal media, which contain amino acids, vitamins, inorganic salts, and a carbon source such as glucose, but these basal media formulations must be further supplemented with serum.
- **Reduced-Serum Media:** Another strategy to reduce the undesired effects of serum in cell culture experiments is to use reduced-serum media. Reduced-serum media are basal media formulations enriched with nutrients and animal-derived factors, which reduce the amount of serum that is needed.
- **Serum-Free Media**: Serum-free media (SFM) circumvents issues with using animal sera by replacing the serum with appropriate nutritional and hormonal formulations. Serum-free media

formulations exist for many primary cultures and cell lines, including recombinant protein producing lines of Chinese Hamster Ovary (CHO), various hybridoma cell lines, the insect lines Sf9 and Sf21 (*Spodopterafrugiperda*), and for cell lines that act as hosts for viral production (e.g., 293, VERO, MDCK, MDBK), and others. One of the major advantages of using serum-free media is the ability to make the medium selective for specific cell types by choosing the appropriate combination of growth factors. Using serum in a medium has a number of disadvantages: the physiological variability, the shelf life and consistency, the quality control, the specificity, the availability, the downstream processing, the possibility of contamination, the growth inhibitors, the standardization and the costs. Using serum-free media and defined media supplements (Nutridoma-CS, Nutridoma-SP and Transferrin) offers three main advantages: The ability to make a medium selective for a particular cell type. The possibility of switching from growth-enhancing medium for propagation to a differentiation-inducing medium. The possibility of bioassays (e.g., protein production) free from interference with serum proteins (easier downstream processing).

- **Media Recommendations:** Many continuous mammalian cell lines can be maintained on a relatively simple medium such as MEM supplemented with serum, and a culture grown in MEM can probably be just as easily grown in DMEM or Medium 199. However, when a specialized function is expressed, a more complex medium may be required. Information for selecting the appropriate medium for a given cell type is usually available in published literature, and may also be obtained from the source of the cells or cell banks. If there is no information available on the appropriate medium for your cell type, choose the growth medium and serum empirically or test several different media for best results. In general, a good place

to start is MEM for adherent cells and RPMI-1640 for suspension cells. The conditions listed below (Table 1) can be used as a guide line when setting up a new mammalian cell culture. Insect cells are cultured in growth media that are usually more acidic that those used for mammalian cells such as TNM-FH and Grace's medium.

Table 1: MammalianCellCulture and medium

CellLine	Cell Type	Species	Tissue	Medium*
293	Fibroblast	Human	Embryonic kidney	MEM and 10% FBS
3T6	Fibroblast	Mouse	Embryo	DMEM, 10% FBS
A549	Epithelial	Human	Lung carcinoma	F-12K, 10% FBS
A9	Fibroblast	Mouse	Connective tissue	DMEM, 10% FBS
AtT-20	Epithelial	Mouse	Pituitary tumor	F-10, 15% horse serum, and 2 .5% FBS
BALB/3T3	Fibroblast	Mouse	Embryo	DMEM, 10% FBS
BHK-21	Fibroblast	Hamster	Kidney	GMEM, 10% FBS, or MEM, 10% FBS and NEAA
BHL-100	Epithelial	Human	Breast	McCoy'5A, 10% FBS
BT	Fibroblast	Bovine	Turbinate cells	MEM, 10% FBS, and NEAA
Caco-2	Epithelial	Human	Colon adeno carcinoma	MEM, 20% FBS, and NEAA
Chang	Epithelial	Human	Liver	BME, 10% calf serum
CHO-K1	Epithelial	Hamster	Ovary	F-12, 10% FBS

Clone 9	Epithelial	Rat	Liver	F-12K, 10% FBS
Clone M-3	Epithelial	Mouse	Melanoma	F-10, 15% horse serum, and 2.5% FBS
COS-1, COS-3, COS-7	Fibroblast	Monkey	Kidney	DMEM, 10% FBS
CRFK	Epithelial	Cat	Kidney	MEM, 10% FBS, and NEAA
CV-1	Fibroblast	Monkey	Kidney	MEM, 10% FBS
D-17	Epithelial	Dog	Osteosarcoma	MEM, 10% FBS, and NEAA
Daudi	Lymphoblast	Human	Blood from a lymphoma patient	RPMI-1640, 10% FBS
GH1, GH3	Epithelial	Rat	Pituitary tumor	F-10, 15% horse serum, and 2.5% FBS
H9	Lymphoblast	Human	T-cell lymphoma	RPMI-1640, 20% FBS
HaK	Epithelial	Hamster	Kidney	BME, 10% calf serum
HCT-15	Epithelial	Human	Colorectal adenocarcinoma	RPMI-1640, 10% FBS
HeLa	Epithelial	Human	Cervix carcinoma	MEM, 10% FBS, and NEAA (in suspension, S-MEM)
HEp-2	Epithelial	Human	Larynx carcinoma	MEM, 10% FBS
HL-60	Lymphoblast	Human	Promyeolocytic leukemia	RPMI-1640, 20% FBS

HT-1080	Epithelial	Human	Fibrosarcoma	MEM, 10% HI FBS, and NEAA
HT-29	Epithelial	Human	Colon adenocarcinoma	McCoy's 5A, 10% FBS
HUVEC	Endothelial	Human	Umbilical cord	F-12K, 10% FBS, and 100 µg/m heparin
I-10	Epithelial	Mouse	Testicular tumor	F-10, 15% horse serum, and 2 .5% FBS
IM-9	Lymphoblast	Human	Marrow (myeloma patient)	RPMI-1640, 10% FBS
JEG-2	Epithelial	Human	Choriocarcinoma	MEM, 10% FBS
Jensen	Fibroblast	Rat	Sarcoma	McCoy's 5A, 5% FBS
Jurkat	Lymphoblast	Human	Lymphoma	RPMI-1640, 10% FBS
K-562	Lymphoblast	Human	Myelogenous leukemia	RPMI-1640, 10% FBS
KB	Epithelial	Human	Oral carcinoma	MEM, 10% FBS, and NEAA
KG-1	Myeloblast	Human	Marrow (erythroleukemia) patient	IMDM, 20% FBS
L2	Epithelial	Rat	Lung	F-12K, 10%FBS
LLC-WRC 256	Epithelial	Rat	Carcinoma	Medium 199, 5% horse serum
McCoy	Fibroblast	Mouse	Unknown	MEM, 10% FBS
MCF7	Epithelial	Human	Breast adenocarcinoma	MEM, 10% FBS, NEAA, and 10 µg/mL insulin
WI-38	Epithelial	Human	Embryonic lung	BME, 10% FBS

WISH	Epithelial	Human	Amnion	BME, 10% FBS
XC	Epithelial	Rat	Sarcoma	MEM, 10% FBS, and NEAA
Y-1	Epithelial	Mouse	Tumor from adrena	F-10, 15% horse serum, and 2.5% FBS

BME: Basal Medium Eagle; **DMEM**: Dulbecco's Modified Eagle Medium; **FBS**: Fetal Bovine Serum; **GMEM**: Glasgow Minimum Essential Medium; **IMDM**: Iscove's Modified Dulbecco's Medium; **MEM**: Minimum Essential Medium; **NEAA**: Non-Essential Amino Acids Solution; **TNM-FH:** *Trichoplusiani* Medium-Formulation Hink (i.e. Grace's Insect Medium, Supplemented)

Cell Culture

Cell culture is one of the major tools used in cellular and molecular biology, providing excellent model systems for studying the normal physiology and biochemistry of cells (e.g., metabolic studies, aging), the effects of drugs and toxic compounds on the cells, and mutagenesis and carcinogenesis. It is also used in drug screening and development, and large scale manufacturing of biological compounds (e.g., vaccines, therapeutic proteins). The major advantage of using cell culture for any of these applications is the consistency and reproducibility of results that can be obtained from using a batch of clonal cells. When the cells are removed from the organ fragments prior to, or during cultivation, thus disrupting their normal relationships with neighboring cells, it is called cell culture.

Tissue culture is the general term for the removal of cells from an animal or plant and their subsequent growth in a favorable artificial environment. The cells may be removed from the tissue directly and disaggregated by enzymatic or mechanical means before cultivation, or they may be derived from a cell line or cell strain that has already been

already established. The culture of whole organs or intact organ fragments with the intent of studying their continued function or development is called organ culture.

Primary Culture:

Primary culture refers to the stage of the culture after the cells are isolated from the tissue and proliferated under the appropriate conditions until they occupy all of the available substrate (i.e., reach confluence). There are two basic methods for doing this.

i. **Explant Cultures**, small pieces of tissue are attached to a glass or treated plastic culture vessel and bathed in culture medium. After a few days, individual cells will move from the tissue explant out on the culture vessel surface or substrate where they will begin to divide and grow.

ii. **Enzymatic Dissociation** more widely used method speeds up this process by adding digesting enzymes, such as trypsin or collagenase, to the tissue fragments to dissolve the cement holding the cells together. This creates a suspension of single cells that are then placed into culture vessels containing culture medium and allowed to grow and divide.

Subculturing:

When the cells in the primary culture vessel have grown and filled up all of the available culture substrate, they must be subcultured (i.e., passaged) by transferring them to a new vessel with fresh growth medium to provide more room for continued growth.

Buying and Borrowing

An alternative to establishing cultures by primary culture is to buy established cell cultures from organization such as the American

Type Culture Collection (ATCC) or the Coriell Institute for Medical Research.

- **Cell Line:** After the first subculture, the primary culture becomes known as a cell line or subclone.Cell lines derived from primary cultures have a limited life span (i.e., they are finite), and as they are passaged, cells with the highest growth capacity predominate, resulting in a degree of genotypic and phenotypic uniformity in the population.
- **Cell Strain:** If a subpopulation of a cell line is positively selected from the culture by cloning or some other method, this cell line becomes a cell strain. A cell strain often acquires additional genetic changes subsequent to the initiation of the parent line.
- **Finite vs. Continuous Cell Lines:** Normal cells usually divide only a limited number of times before losing their ability to proliferate, which is a genetically determined event known as senescence; these cell lines are known as **Finite**. However, some cell lines become immortal through a process called transformation, which can occur spontaneously or can be chemically or virally induced. When a finite cell line undergoes transformation and acquires the ability to divideindefinitely, it becomes a **Continuous** cell line.

There are two basic systems for growing cells in culture, as monolayers on an artificial substrate (i.e., **adherent culture**) or free-floating in the culture medium (**suspensionculture**). The majority of the cells derived from vertebrates, with the exception of hematopoietic cell lines and a few others are anchorage-dependent and have to be cultured on a suitable substrate that is specifically treated to allow cell adhesion and spreading (i.e., tissue-culture treated). However, many cell lines can also be adapted for suspension culture. Similarly, most of the commercially available insect cell lines grow well in monolayer or

suspension culture. Cells that are cultured in suspension can be maintained in culture flasks that are not tissue-culture treated, but as the culture volume to surface area is increased beyond which adequate gas exchange is hindered (usually 0.2–0.5 mL/cm2), the medium requires agitation. This agitation is usually achieved with a magnetic stirrer or rotating spinner flasks.

AdherentCulture	SuspensionCulture
Appropriate for most cell types, including primary cultures.	Appropriate for cells adapted to suspension culture and a few other cell lines that are nonadhesive (e.g., hematopoietic)
Requires periodic passaging, but allows easy visual inspection under inverted microscope.	Easier to passage, but requires daily cell counts and viability determination to follow growth patterns; culture can be diluted to stimulate growth.
Cells are dissociated enzymatically (e.g., TrypLE™ Express, trypsin) or mechanically.	Does not require enzymatic or mechanical dissociation.
Growth is limited by surface area, which may limit product yields.	Growth is limited by concentration of cells in the medium, which allows easy scale-up.
Requires tissue-culture treated vessel.	Can be maintained in culture vessels that are not tissue-culture treated, but requires agitation (i.e., shaking or stirring) for adequate gas exchange.
Used for cytology, harvesting products continuously, and many research applications.	Used for bulk protein production, batch harvesting, and many research applications

Maintenance:

Once a culture is initiated, whether it is a primary culture or a subculture, it will need periodic medium changes. For example, HeLa cells are usually subcultured once per week. Other cell lines may be subcultured only every two, three or even four weeks.

Modification of Cell Morphology:

Prior to use, cells should always be checked for any signs of deterioration, such as granularity around the nucleus, cytoplasmic vacuolation, or rounding of the cells with detachment from substrate. Such signs may imply that the culture requires a medium change or may indicate a more serious problem (inadequate or toxic serum/medium, microbial contamination or senescence of the cell line).

Replacement of the Medium:

Four factors indicate the need for the replacement of culture medium,

1. **Drop in pH**: Most cells stop growing as the pH falls from pH7.0 to pH 6.5 and start to lose viability between pH 6.5 and pH 6.0. *(As the pH drops, the indicator in the mediumchanges from red through orange to yellow*.)
2. **Cell Concentrations**: High cell concentrations exhaust the medium faster than low concentrations.
3. **Cell Type**: Normal cells usually stop dividing at high density due to cell crowding, growth factor depletion, etc. The cells arrest in the G1 phase of the cell cycle and deteriorate very little, even if left for two to three weeks (or longer).
4. **Deterioration of Morphology**: This factor should be checked frequently. You should always be aware of the morphology since this may reveal the presence of contamination.

Criteria for Subculture

Density of the Culture:

Cells should be subcultured prior to confluence. The ideal method for determining the correct seedingdensity is to perform a growth curve at different seeding concentrations. This allows you to determine the minimum concentration that will give a short lag period and early entry into rapid logarithmic growth.

Exhaustion of Medium:

Medium requires periodic replacement. If the pH falls too rapidly, subculture may be required. Time since Last Subculture orRoutine subculture is best performed according to a strict schedule, so that reproducible behavior is achieved. It is essential to become familiar with the growth cell cycle for each cell line. Cells at different phases behave differently with respect to proliferation, enzyme activity, glycolysis and respiration, synthesis of specialized products, etc.

Requirements for Other Procedures:

When cells require operations other than routine propagation (e.g., increasing stock, changing vessel or medium), this procedure should ideally be done at the regular subculture time. Cells should not be subcultured while still in the lag phase; cells should always be taken between the middle of the log phase and the plateau phase as determined during a previous subculture Fig. 1 (unless experimental requirements dictate different timing)

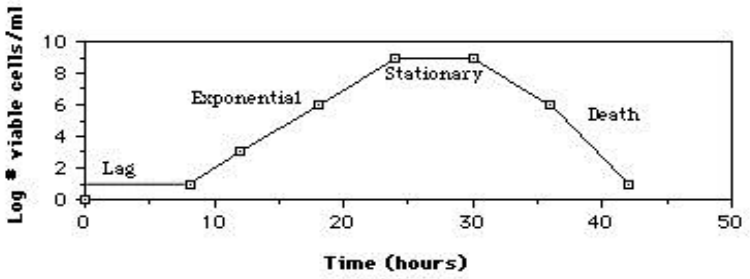


Figure 1: Growthcurve of cell culture

Culture Conditions

Culture conditions vary widely for each cell type, but the artificial environment in which the cells are cultured invariably consists of a suitable vessel containing a substrate or medium that supplies the essential nutrients (amino acids, carbohydrates, vitamins, minerals), growth factors, hormones, and gases (O2, CO2), and regulates the physicochemical environment (pH, osmotic pressure, temperature).

Mammalian Cell:

Morphology Most mammalian cells in culture can be divided in to three basic categories based on theirmorphology (Fig. 2)

1) Fibroblastic (or fibroblast-like) cells are bipolar or multipolar and have elongated shapes. They grow attached to a substrate.
2) Epithelial-like cells are polygonal in shape with more regular dimensions, and grow attached to a substrate in discrete patches.
3) Lymphoblast-like cells are spherical in shape and they are usually grown in suspension without attaching to a surface.

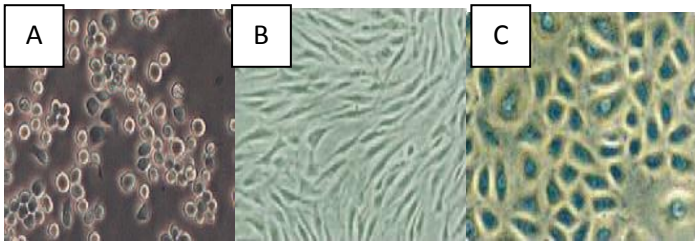


Figure 2: Mammalian cell line (A) Fibroblast (B) Epithelial and (C) Lymphoblast

In addition to the basic categories listed above, certain cells display morphological characteristics specific to their specialized role in host.

Aseptic Techniques

To minimize the risk of contamination, follow these 5 rules:

1. Always check the cells carefully before handling (by eye and on a microscope). Become familiar with the indicators of abnormal cell growth.
2. Whenever possible, maintain cultures without antibiotics for at least part of the time, to reveal cryptic contamination.
3. Check sterility of all reagents before use.
4. Use dedicated media and reagents; do not share with other cell lines.
5. Maintain a high standard of sterility at all steps.

Mycoplasma contamination, which may slow cell growth, cannot be checked under a regular microscope. To confirm or rule out such contamination, use a mycoplasma test (e.g. Roche Applied Science Mycoplasma PCR ELISA Kit).

Environment:

There should be a laminar flow hood in the room dedicated to cell culture, and this hood should be usedfor all culture manipulations and storage of all equipment. The hood must be placed away from traffic orequipment that might generate air currents (e.g., centrifuges, refrigerators and freezers).Always carefully clean the hood before and after your procedure. Remove all unneeded items.It is crucial to always keep the work surface clean and tidy. To achieve this, follow these rules:

- Use 80% ethanol to clean the surface before starting.
- Place and keep on this surface only the items required for your procedure. This will reduce the possibility of contact between sterile and non-sterile items and facilitate culture manipulations.
- Clear space in the center of the bench, not just the front edge.
- Avoid spills, if they happen immediately clean the area.
- Remove everything when you are done, and again clean the work surface.

Reagents and media obtained from commercial suppliers will already have undergone strict quality testing. Most of the bottles are wrapped in polyethylene. The wrapping should be removed outside the hood. Unwrapped bottles should be cleaned with 80% ethanol whenever they are removed from the refrigerator or from a water bath. Regularly clean the refrigerator, the incubator and the water bath to avoid growth of mold or fungi. Imported cell lines should always be quarantined before being incorporated into your main stock. Do not perpetually use antibiotics; they will suppress some contaminants, but will not eliminate them.

Handling:

Special care should be taken with caps. Use deep screw caps in preference to stoppers. When working on an open bench, flame glass

pipettes and necks of the bottles before and after each use. Always use the pipettes which are best adapted your procedure; regularly clean them and check their calibration. Use a multi-channel pipette instead of a single pipette if you are working with multiwell plates. This will reduce both the time required to perform the procedure and the probability of contamination. Prepare as many reagents and equipment as possible in advance, to reduce the time the cultures are kept out of the incubator.

Cell Lines (contamination, cryopreservation):

Cryopreservation Cell lines in continuous culture are prone to genetic drift, finite cell lines are fated forsenescence, all cell cultures are susceptible to microbial contamination, and even thebest-run laboratories can experience equipment failure. Because an established cell lineis a valuable resource and its replacement is expensive and time consuming, it is vitallyimportant that they are frozen down and preserved for long-term storage.As soon as a small surplus of cells becomes available from subculturing, they should befrozen as a seed stock, protected, and not be made available for general laboratory use.Working stocks can be prepared and replenished from frozen seed stocks. If the seedstocks become depleted, cryopreserved working stocks can then serve as a source forpreparing a fresh seed stock with a minimum increase in generation number from theinitial freezing.The best method for cryopreserving cultured cells is storing them in liquid nitrogen incomplete medium in the presence of a cryoprotective agent such as dimethyl sulfoxide(DMSO). Cryoprotective agents reduce the freezing point of the medium and also allow aslower cooling rate, greatly reducing the risk of ice crystal formation, which can damagecells and cause cell death.

Note:

DMSO is known to facilitate the entry of organic molecules into tissues. Handle reagents containing DMSO using equipment and practices appropriate for the hazards posed by such materials. Dispose of the reagents in compliance with local regulations.

Guidelines for Cryopreservation Following the guidelines below are essential for cryopreserving your cell lines for future use. As with other cell culture procedures, we recommend that you closely follow the instructions provided with your cell line for best results.

- Freeze your cultured cells at a high concentration and at as low a passage number as possible. Make sure that the cells are at least 90% viable before freezing. Note that the optimal freezing conditions depend on the cell line in use.
- Freeze the cells slowly by reducing the temperature at approximately 1^{o}C per minute using a controlled rate cryo-freezer or a cryo-freezing container such as “Mr. Frosty,” available from NALGENE labware (Nalgene Nunc)
- Always use the recommended freezing medium. The freezing medium should contain a cryoprotective agent such as DMSO or glycerol.
- Store the frozen cells below –70^{o}C; frozen cells begin to deteriorate above –50^{o}C.
- Always use sterile cryovials for storing frozen cells. Cryovials containing the frozen cells may be stored immersed in liquid nitrogen or in the gas phase above the liquid nitrogen.
- Always wear personal protective equipment.
- All solutions and equipment that come in contact with the cells must be sterile. Always use proper sterile technique and work in a laminar flow hood.

- **Freezing Medium**:

Always use the recommended freezing medium for cryopreserving your cells. The freezing medium should contain a cryoprotective agent such as DMSO or glycerol.

Cryopreservation Medium

- **Recovery™**:

Cell Culture Freezing Medium is a ready-to-use complete cryopreservation medium for mammalian cell cultures, containing an optimized ratio of fetal bovine serum to bovine serum for improved cell viability and cell recovery after thawing.

- **Synth-a-Freeze**:

Cryopreservation Medium is a chemically defined, protein free, sterile cryopreservation medium containing 10% DMSO that is suitable for the cryopreservation of many stem and primary cell types with the exception of melanocytes.

Protocol for Cryopreserving Cultured Cells

The following protocol describes a general procedure for cryopreserving cultured cells. For detailed protocols, always refer to the cell-specific product insert.

1. Prepare freezing medium and store at 2^{o}C to 8^{o}C until use. Note that the appropriate freezing medium depends on the cell line.
2. For adherent cells, gently detach cells from the tissue culture vessel following the procedure used during the subculture. Resuspend the cells in complete medium required for that cell type.
3. Determine the total number of cells and percent viability using a hemacytometer, cell counter and Trypan Blue exclusion, or the Countess, Automated Cell Counter. According to the desired

viable cell density, calculate the required volume of freezing medium.

4. Centrifuge the cell suspension at approximately 100–200 g for 5 to 10 minutes; aseptically decant supernatant without disturbing the cell pellet.
1. Note: Centrifugation speed and duration varies depending on the cell type.
5. Resuspend the cell pellet in cold freezing medium at the recommended viable cell density for the specific cell type.
6. Dispense aliquots of the cell suspension into cryogenic storage vials. As you aliquot them, frequently and gently mix the cells to maintain a homogeneous cell suspension.
7. Freeze the cells in a controlled rate freezing apparatus, decreasing the temperature approximately $1^{o}C$ per minute. Alternatively, place the cryovials containing the cells in an isopropanol chamber and store them at $–80^{o}C$ overnight.
8. Transfer frozen cells to liquid nitrogen, and store them in the gas phase above the liquid nitrogen.

Thawing Frozen Cells

Protocol for Thawing Frozen Cells

The following protocol describes a general procedure for thawing cryopreserved cells. For detailed protocols, always refer to the cell-specific product insert.

1. Remove the cryovial containing the frozen cells from liquid nitrogen storage and immediately place it into a $37^{o}C$ water bath.
2. Quickly thaw the cells (< 1 minute) by gently swirling the vial in the $37^{o}C$ water bath until there is just a small bit of ice left in the vial.

3. Transfer the vial it into a laminar flow hood. Before opening, wipe the outside of the vial with 70% ethanol.
4. Transfer the thawed cells drop wise into the centrifuge tube containing the desired amount of pre-warmed complete growth medium appropriate for your cell line.
5. Centrifuge the cell suspension at approximately 200~ g for 5–10 minutes. The actual centrifugation speed and duration varies depending on the cell type.
6. After the centrifugation, check the clarity of supernatant and visibility of a complete pellet. Aseptically decant the supernatant without disturbing the cell pellet.
7. Gently re-suspend the cells in complete growth medium, and transfer them into the appropriate culture vessel and into the recommended culture environment.

Note: The appropriate flask size depends on the number of cells frozen in the cryovial, and the culture environment varies based on the cell and media type.

Biological Contamination

Contamination of cell cultures is the common problem encountered in cell culture laboratories, sometimes with very serious consequences. Cell culture contaminants can be divided into two main categories, **chemical contaminants** such as impurities in media, sera, and water, endotoxins, plasticizers, and detergents, and **biological contaminants** such as bacteria, molds, yeasts, viruses, mycoplasma, as well as cross contamination by other cell lines. While it is impossible to eliminate contamination entirely, it is possible to reduce its frequency and seriousness by gaining a thorough understanding of their sources and by following good aseptic technique. This section provides an overview of major types of biological contamination. Bacterial

contamination is easily detected by visual inspection of the culture within a few days of it becoming infected; infected cultures usually appear cloudy (i.e., turbid), sometimes with a thin film on the surface. Sudden drops in the pH of the culture medium are also frequently encountered. Under a low power microscope, the bacteria appear as tiny, moving granules between the cells, and observation under a high-power microscope can resolve the shapes of individual bacteria. The simulated images below show an adherent 293 cell culture contaminated with *E. coli* (Fig. 3).

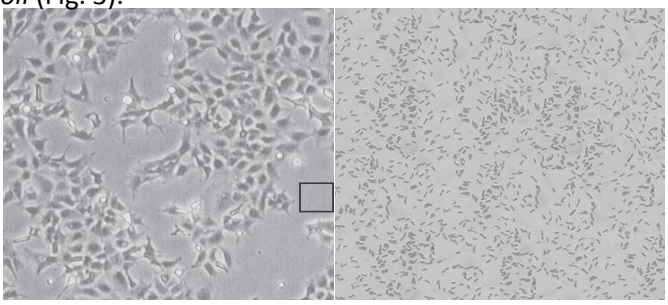


Figure 3: Simulated phase contrast images of adherent 293 cells contaminated with *E. coli*. The spaces between the adherent cells show tiny, shimmering granules under low power microscopy, but the individual bacteria are not easily distinguishable (panel A). Further magnification of the area enclosed by the black square resolves the individual *E. coli* cells, which are typically rod-shaped and are about 2 µm long and 0.5 µm in diameter. Each side of the black square in panel A is 100 µm.

Yeasts are unicellular eukaryotic microorganisms in the kingdom of Fungi, ranging in size from a few micrometers (typically) up to 40 micrometers (rarely). Like bacterial contamination, cultures contaminated with yeasts become turbid, especially if the contamination is in an advanced stage. There is very little change in the pH of the culture contaminated by yeasts until the contamination becomes heavy,

at which stage the pH usually increases. Under microscopy, yeast appears as individual ovoid or spherical particles, which may bud off smaller particles. The simulated image below shows adherent 293 cell culture 24 hours after plating that is infected with yeast (Fig. 4).

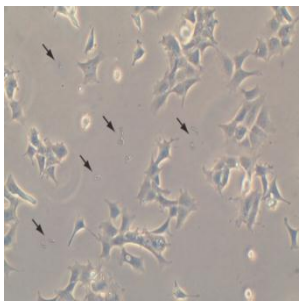


Figure 4: Simulated phase contrast images of 293 cells in adherent culture that is contaminated with yeast. The contaminating yeast cells appear as ovoid particles, budding off smaller particles as they replicate.

Similar to yeast contamination, the pH of the culture remains stable in the initial stages of contamination, then rapidly increases as the culture become more heavily infected and becomes turbid. Under microscopy, the mycelia usually appear as thin, wisp-like filaments, and sometimes as denser clumps of spores. Spores of many mold species can survive extremely harsh and inhospitable environments in their dormant stage, only to become activated when they encounter suitable growth conditions. Viruses are microscopic infectious agents that take over the host cells machinery to reproduce. Their extremely small size makes them very difficult to detect in culture, and to remove them from reagents used in cell culture laboratories. Because most viruses have very stringent requirements for their host, they usually do not adversely affect cell cultures from species other than their host. However, using virally infected cell cultures can present a serious health hazard to the

laboratory personnel, especially if human or primate cells are cultured in the laboratory. Viral infection of cell cultures can be detected by electron microscopy, immunostaining with a panel of antibodies, ELISA assays, or PCR with appropriate viral primers.

Protocol for Microbial Decontamination

1. Collect the contaminated medium carefully. If possible, the organism should be tested for sensitivity to a range of individual antibiotics. If not, autoclave the medium or add hypochlorite.
2. Washthe cells in DBSS (Hanks BSS without bicarbonate, with Penicillin, Streptomycin, Amphotericin B and Kanamycin or Gentamycin). For monolayers, rinse the culture 3 times with DBSS, trypsinize and then wash the cells twice more in DBSS by centrifugation and re-suspension. For suspension cultures, wash the culture five times (in DBSS) by centrifugation and re-suspension.
3. Reseed a fresh flask at the lowest reasonable seeding density, depending on cell type.
4. Add high antibiotic medium and change the culture every 2 days.
5. Subculture in a high antibiotic medium. Repeat Steps 1 to 4for three subcultures.
6. Remove the antibiotics, and culture the cells without them for a further three subcultures.
7. Recheck the cultures (phase contrast microscopy, Hoechst staining).
8. Culture the cells for a further two months without antibiotics, and check to make sure that all contamination has been eliminated.

Mycoplasmas are simple bacteria that lack a cell wall, and they are considered the smallest self-replicating organism. Because of their extremely small size (typically less than one micrometer), mycoplasma are very difficult to detect until they achieve extremely high densities and cause the cell culture to deteriorate; until then, there are often no visible signs of infection. Chronic mycoplasma infections might manifest themselves with decreased rate of cell proliferation, reduced saturation density, and agglutination in suspension cultures; (Fig. 5) however, the only assured way of detecting mycoplasma contamination is by testing the cultures periodically using fluorescent staining (e.g., Hoechst 33258), ELISA, PCR, immunostaining, autoradiography, or microbiological assays.

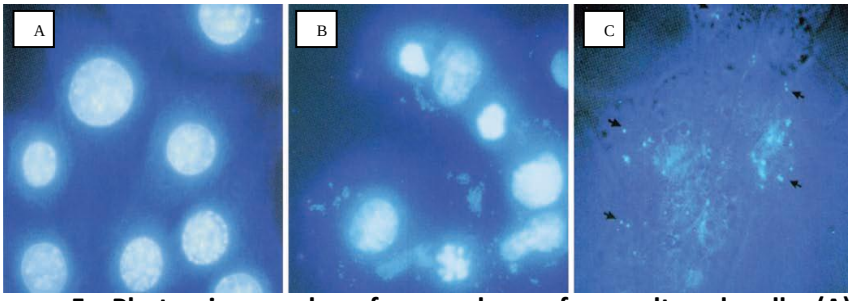


Figure 5: Photomicrographs of mycoplasma-free cultured cells (A) infected with mycoplasma (B and C).

Protocol for Treating Mycoplasma-contaminated Cell Cultures with BM Cyclin

Remove culture medium from culture vessels by aspiration. Add new culture medium containing BM Cyclin 1 (4 µl of stock solution/ml, final concentration 10 µg/ml). Cultivate the cells for 3 days as usual. Remove culture medium, add new culture medium containing BM Cyclin 2 (4 µl of stock solution/ml, final concentration 5 µg/ml). Cultivate the cells for 4 days, repeat the above cycle twice. Cross-Contamination While not as common as microbial contamination, extensive cross-

contamination of many cell lines with HeLa and other fast growing cell lines is a clearly-established problem with serious consequences. Obtaining cell lines from reputable cell banks, periodically checking the characteristics of the cell lines, and practicing good aseptic technique are practices that will help you avoid cross-contamination. DNA fingerprinting, karyotype analysis, and isotype analysis can confirm the presence or absence of cross contamination in your cell cultures. Using Antibiotics Antibiotics should not be used routinely in cell culture, because their continuous use encourages the development of antibiotic resistant strains and allows low-level contamination to persist, which can develop into full-scale contamination once the antibiotic is removed from media, and may hide mycoplasma infections and other cryptic contaminants. Further, some antibiotics might cross react with the cells and interfere with the cellular processes under investigation. Antibiotics should only be used as a last resort and only for short term applications, and they should be removed from the culture as soon as possible. If they are used in the long term, antibiotic-free cultures should be maintained in parallel as a control for cryptic infections.

Cell Viability (quantitation and cytotoxicity):

The measurement of cell viability plays a fundamental role in all forms of cell culture. Sometimes it is the main purpose of the experiment, such as in toxicity assays. Alternatively, cell viability can be used to correlate cell behaviour to cell number, providing a more accurate picture, for example anabolic activity. There are wide arrays of cell viability methods which range from the most routine trypan blue dye exclusion assay to highly complex analysis of individual cells, such as using RAMAN microscopy. The cost, speed, and complexity of equipment required will all play a role in determining the assay used. This chapter provides an overview of many of the assays available today. Cell viability

is a determination of living or dead cells, based on a total cell sample. Viability measurements may be used to evaluate the death or life of cancerous cells and the rejection of implanted organs. In other applications, these tests might calculate the effectiveness of a pesticide or insecticide, or evaluate environmental damage due to toxins.

Factors to Consider When Choosing a Cell Viability Assay:

Among the many factors to consider when choosing a cell-based assay, the primary concern for many researchers is the ease of use. Homogeneous assays do not require removal of culture medium, cell washes or centrifugation steps. When choosing an assay, the time required for reagent preparation and the total length of time necessary to develop a signal from the assay chemistry should be considered. The stability of the absorbance, fluorescence or luminescence signal is another important factor that provides convenience and flexibility in recording data and minimizes differences when processing large batches of plates. Another factor to consider when selecting an assay is sensitivity of detection. Detection sensitivity will vary with cell type if you choose to measure a metabolic marker, such as ATP level or MTS tetrazolium reduction. The signal-to-background ratios of some assays may be improved by increasing incubation time. The sensitivity not only depends upon the parameter being measured but also on other parameters of the model system such as the plate format and number of cells used per well. Cytotoxicity assays that are designed to detect a change in viability in a population of 10,000 cells may not require the most sensitive assay technology. For example, a tetrazolium assay should easily detect the difference between 10,000 and 8,000 viable cells. On the other hand, assay model systems that use low cell numbers in a high-density multiwell plate format may require maximum sensitivity of detection such as that achieved with the luminescent ATP assay

technology.For researchers using automated screening systems, the reagent stability and compatibility with robotic components is often a concern. The assay reagents must be stable at ambient temperature for an adequate period of time to complete dispensing into several plates. In addition, the signal generated by the assay should also be stable for extended periods of time to allow flexibility for recording data. For example, the luminescent signal from the ATP assay has a half-life of about 5 hours, providing adequate flexibility. With other formats such as the MTS tetrazolium assay or the LDH release assay, the signal can be stabilized by the addition of a detergent-containing stop solution. In some cases the choice of assay may be dictated by the availability of instrumentation to detect absorbance, fluorescence or luminescence. The Promega portfolio of products contains an optional detection format for each of the three major classes of cell-based assays (viability, cytotoxicity or apoptosis). In addition, results from some assays such as the ATP assay can be recorded with more than one type of instrument (luminometer, fluorometer or CCD camera).

Cost is an important consideration for every researcher; however, many factors that influence the total cost of running an assayare often overlooked. All of the assays described above are homogeneous and as such are more efficient than multistep assays. For example, even though the reagent cost of an ATP assay may be higher than other assays, the speed (time savings), sensitivity (cell sample savings) and accuracy may outweigh the initial cost. Assays with good detection sensitivity that are easier to scale down to 384 or 1536well formats may result in savings of cell culture reagents and enable testing of very small quantities of expensive or rare test compounds. The ability to gather more than one set of data from the same sample (i.e., multiplexing) also may contribute to saving time and effort. Multiplexing more than one assay in the same culture well can provide internal

controls and eliminate the need to repeat work. For instance, the LDH-release assay is an example of an assay that can be multiplexed. The LDH-release assay offers the opportunity to gather cytotoxicity data from small aliquots of culture supernatant that can be removed to a separate assay plate, thus leaving the original assay plate available for any other assay such as gene reporter analysis, image analysis, etc. Several of our homogeneous apoptosis and viability assays can be multiplexed without transferring media, allowing researchers to assay multiple parameters in the same sample well.

Reproducibility of data is an important consideration when choosing a commercial assay. However, for most cell-based assays, the variation among replicate samples is more likely to be caused by the cells rather than the assay chemistry. Variations during plating of cells can be magnified by using cells lines that tend to form clumps rather than a suspension of individual cells. Extended incubation periods and edge effects in plates may also lead to decreased reproducibility among replicates and less desirable Z'-factor values.

Cell Viability Assays that Measure ATP

CellTiter-Glo® Luminescent Cell Viability Assay

The CellTiter-Glo® Luminescent Cell Viability Assay is a homogeneous method to determine the number of viable cells in culture. Detection is based on using the luciferase reaction to measure the amount of ATP from viable cells. The amount of ATP in cells correlates with cell viability. Within minutes after a loss of membrane integrity, cells lose the ability to synthesize ATP, and endogenous ATPases destroy any remaining ATP; thus the levels of ATP fall precipitously. The CellTiter-Glo® Reagent does three things upon addition to cells. It lyses cell membranes to release ATP; it inhibits endogenous ATPases, and it provides luciferin, luciferase and other reagents

necessary to measure ATP using a bioluminescent reaction. The unique properties of a proprietary stable luciferase mutant enabled a robust, single-addition reagent. The "glow-type" signal can be recorded with a luminometer, CCD camera or modified fluorometer and generally has a half-life of five hours, providing a consistent signal across large batches of plates. The CellTiter-Glo® Assay is extremely sensitive and can detect as few as 10 cells. The luminescent signal can be detected as soon as 10 minutes after adding reagent or several hours' later, providing flexibility for batch processing of plates.

Cell Viability Assays that Measure Metabolic Capacity

CellTiter-Blue® Cell Viability Assay (resazurin)

The CellTiter-Blue® Cell Viability Assay uses an optimized reagent containing resazurin. The homogeneous procedure involves adding the reagent directly to cells in culture at a recommended ratio of 20µl of reagent to 100µl of culture medium. The assay plates are incubated at 37°C for 1–4 hours to allow viable cells to convert resazurin to the fluorescent resorufin product. The conversion of resazurin to fluorescent resorufin is proportional to the number of metabolically active, viable cells present in a population. The signal is recorded using a standard multiwellfluorometer. Because different cell types have different abilities to reduce resazurin, optimizing the length of incubation with the CellTiter-Blue® Reagent can improve assay sensitivity for a given model system. The detection sensitivity is intermediate between the ATP assay and the MTS reduction assay.

Cytotoxicity Assays: Determining the Number of Live and Dead Cells in a Cell Population, MultiTox-Fluor Multiplex Cytotoxicity Assay:

Cell-based assays are important tools for contemporary biology and drug discovery because of their predictive potential for in vivo

applications. However, the same cellular complexity that allows the study of regulatory elements, signaling cascades or test compound bio-kinetic profiles also can complicate data interpretation by inherent biological variation. Therefore, researchers often need to normalize assay responses to cell viability after experimental manipulation. Although assays for determining cell viability and cytotoxicity that are based on ATP, reduction potential and LDH release are useful and cost-effective methods, they have limits in the types of multiplexed assays that can be performed along with them. The MultiTox-Fluor Multiplex Cytotoxicity Assay (Cat. No. G9200, G9201, G9202) is a homogeneous, single-reagent-addition format that allows the measurement of the relative number of live and dead cells in a cell population. This assay gives ratiometric, inversely proportional values of viability and cytotoxicity that are useful for normalizing data to cell number. Also, this reagent is compatible with additional fluorescent and luminescent chemistries.

- **Assays to Detect Apoptosis:**

A variety of methods are available for detecting apoptosis to determine the mechanism of cell death. The Caspase-Glo® Assays are highly sensitive, luminescent assays with a simple “add, mix, measure” protocol that can be used to detect caspase-8, caspase-9 and caspase-3/7 activities. If you prefer a fluorescent assay, the Apo-ONE® Homogeneous Caspase-3/7 Assay is useful and, like the Caspase-Glo®Assays, can be multiplexed with other assays. A later marker of apoptosis is TUNEL analysis to identify the presence of oligonucleosomal DNA fragments in cells. The DeadEnd™ Fluorometric and the DeadEnd™ Colorimetric TUNEL Assays allow users to end-label the DNA fragments to detect apoptosis

- **Cell Counter:**

A cell counter is essential for quantitative growth kinetics, and a great advantage when more than two or three cell lines are cultured in the laboratory. The Countess; Automated Cell Counter is a bench-top instrument designed to measure cell count and viability (live, dead, and total cells) accurately and precisely in less than a minute per sample, using the standard Trypan Blue uptake technique. Using the same amount of sample that you currently use with the hemacytometer, the countess. Automated Cell Counter takes less than a minute per sample for a typical cell count and is compatible with a wide variety of eukaryotic cells.

- **Multiplexing Cell Viability Assays:**

The latest generation of cell-based assays includes luminescent and fluorescent chemistries to measure markers of cell viability, cytotoxicity and apoptosis, as well as to perform reporter analysis. Using these tools researchers can investigate how cells respond to growth factors, cytokines, hormones, mitogens, radiation, effectors, compound libraries and other signaling molecules. However, researchers often need more than one type of data from a sample, so the ability to multiplex, or analyze more than one parameter from a single sample, is desirable.

- **Counting Cells in a Hemacytometer**:

Hemacytometers may be obtained from most major laboratory suppliers (e.g., Baxter Scientific). The procedure below provides some general directions on how to use the hemacytometer.

1. Clean the chamber and cover slip with alcohol. Dry and fix the coverslip in position.
2. Harvest the cells. Add 10 µL of the cells to the hemacytometer. Do not overfill.
3. Place the chamber in the inverted microscope under a 10X objective. Use phase contrast to distinguish the cells.

4. Count the cells in the large, central gridded square (1 mm^2). The gridded square is circled in the graphic below. Multiply by 10^4 to estimate the number of cells per mL. Prepare duplicate samples and average the count.

Trypan Blue:

Exclusion The following procedure will enable you to accurately determine the cell viability. Cell viability is calculated as the number of viable cells divided by the total number of cells within the grids on the hemacytometer. If cells take up trypan blue, they are considered non-viable.

1. Determine the cell density of your cell line suspension using a hemacytometer.
2. Prepare a 0.4% solution of trypan blue in buffered isotonic salt solution, pH 7.2 to 7.3 (i.e., phosphate-buffered saline).
3. Add 0.1 mL of trypan blue stock solution to 1 mL of cells.
4. Load a hemacytometer and examine immediately under a microscope at low magnification.
5. Count the number of blue staining cells and the number of total cells. Cell viability should be at least 95% for healthy log-phase cultures. Remember to correct for the dilution factor.

To calculate the number of viable cells per mL of culture, use the formula below.

Live cell count/ Total cell count =Viability

Determine total viable cell yield using the formula below.

Viable cell count/ Quadrants counted x Dilution factor x Hemocytometer factor x Current volume (mL) = Viable cell yield.

Concentrating Cells:

To concentrate cells from a suspension culture (or resuspended cells from monolayer culture):

1. Transfer the cell suspension to a sterile centrifuge tube of appropriate size andcentrifuge for 10 minutes at 800 ~ g.
   Note: Certain cell lines are very sensitive to centrifugal force.
2. Carefully remove the supernatant without disturbing the cell pellet.
3. Add the desired volume of fresh medium gently to the side of the tube and slowly pipette up and down 2 to 3 times to resuspend the cell pellet.
4. Transfer the cells to the desired, sterile container.

Cell Proliferation:

An alternative way to determine the health of a culture is to perform a cell proliferation assay, i.e. to determine the number of dividing cells. One way of measuring this parameter is by performing clonogenic assays. In these assays, a defined number of cells are plated onto an appropriate matrix and the numbers of colonies that form are counted after a period of growth. Drawbacks to this type of assay are that it is tedious and it is not practical for large numbers of samples. Another way to analyze cell proliferation is to measure DNA Synthesis. In these assays, labeled DNA precursors (4H-thymidine or bromodeoxy-uridine, BrdU (e.g., Roche Applied Science Cell Proliferation ELISA, BrdU (chemiluminescent) Kit) are added to cells and their incorporation into DNA is quantified after incubation. The amount of labeled precursor incorporated into DNA is quantified either by measuring the total amount of labeled DNA in a population, or by detecting the labeled nucleimicroscopically. Cell proliferation can also be measured using more indirect parameters. In these techniques, molecules that regulate the Cell Cycle (also called proliferation markers) are measured either by their activity (e.g., CDK kinase assays) or by quantifying their amounts (e.g., Western blots, ELISA, or immunohistochemistry).

Cell Cycle:

The cell cycle is made up of four phases (Fig. 6). In the M phase (M=mitosis), the chromatin condenses into chromosomes, and the two individual chromatids, which make up the chromosome, segregate to each daughter cell. In the G1 (Gap 1) phase, the cell either progresses toward DNA synthesis or another division cycle or exits the cell cycle reversibly (G0) or irreversibly to commit to differentiation. During G1, the cell is particularly susceptible to control of cell cycle progression; this may occur at a number of restriction points, which determine whether the cell will re-enter the cycle, withdraw from it, or withdraw and differentiate. G1 is followed by the S phase (DNA synthesis), in which the DNA replicates. S in turn is followed by the G2 (Gap 2) phase in which the cell prepares for reentry into mitosis. Checkpoints, at the beginning of DNA synthesis and in G2, determine the integrity of the DNA and will halt the cell cycle to allow either DNA repair or entry into apoptosis if repair is impossible. The Phospho Histone H3 Imaging Kit (Roche) is a convenient method for fast cell cycle analysis by quantification of mitotic cells. Apoptosis, or programmed cell death, is a regulated physiological process whereby a cell can be removed from a population. Characterized by DNA fragmentation, nuclear blebbing, and cell shrinkage, apoptosis can be detected via a number of marker enzymes and kits (see Roche Applied Science products). Roche DNA Fragmentation Imaging Kit is a TUNEL assay-based method for accurate and fast quantitative fluorescence detection of apoptosis in medium to high throughput cellular workflows.

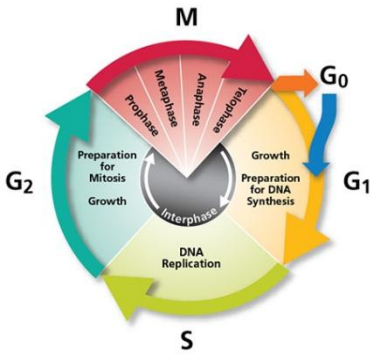


Figure 6: Cell cycle phases

Cytotoxicity:

Cell viability and toxic effects can be assayed using Roche s easy-to-apply one-step Cell Viability Imaging Kit. The indicators of cytotoxicity can vary, depending on the study performed (e.g., Roche Applied Science Cytotoxicity Detection Kit Plus, (LDH)). The cytotoxicity effect can lead to the death of the cells or just to an alteration of their metabolism. This toxic effect can be initiated by addition of compounds or by addition of effector cells. Demonstrating the lack of toxicity of a given compound may require subtle analysis of its interaction with specific targets, e.g. a study of its ability to alter cell signaling or to initiate cell interactions that would give rise to an inflammatory or allergic response.

To test the potential cytotoxicity of compounds/cells, consider the following parameters:

1. **Concentration of Compound**: A wide range of concentrations should be tested to determine the survival curve.
2. **Medium/Serum**: In some cases, the serum may have a masking effect and lead to an underestimation of the cytotoxicity effect.
3. **Duration of the Exposure**: The action of one compound can happen

within a few seconds or over several hours. Cell Density For most of the assays, confluent cells is not used. However, if you want to study the endothelial barrier function, you will need confluent cells in order to see an effect.

4. **Colony Size:** Some agents are cytostatic, i.e. they inhibit cell proliferation but are not cytotoxic. During continuous exposure they may reduce the size of colonies without reducing the number of colonies. In this case, the size of the colonies should be determined by densitometry, automatic colony counting or counting the number of cells per colony with the naked eye.
5. **Solvents**: Some agents to be tested have low solubilities in aqueous media, and it may be necessary to use an organic solvent to dissolve them. Ethanol, propylene glycol and dimethyl sulfoxide have been used for this purpose, but may themselves be toxic to cells. The final concentration of solvent should be maintained as low as possible (<0.5 %) and a solvent control must always be included in the study. Be aware that some organic solvents are not compatible with plastics.
6. The Dose-response relationship describes the biological effect induced by different concentrations of a substance (Fig. 7). This curve should be determined whenever a new study is initiated, in order to fix the optimal conditions for the assay.

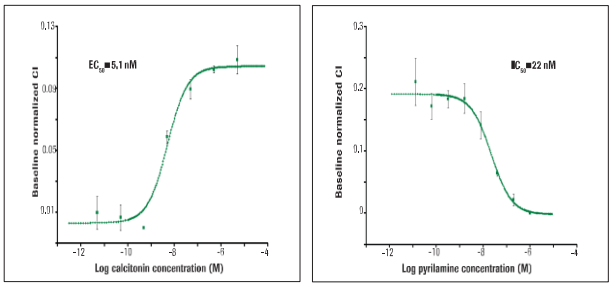


Figure 7: Dose-response curves.

The half-maximal effective concentration, or EC50, refers to the concentration of a compound which induces a response halfway between the baseline and the maximum. The EC50 represents the concentration of a compound where 50% of its maximal effect is observed. The half-maximal inhibitory concentration, or IC50, is the concentration of a compound required to inhibit a process by half. IC50 represents the concentration of a compound that is required for 50% inhibition *in vitro*. The median lethal dose, LD50 (abbreviation for "Lethal Dose, 50%" or LCt50 (Lethal Concentration & Time) of a toxic compound is the dose required to kill half the tested population.

Applications of cell culture

Cell culture is one of the major tools used in cellular and molecular biology, providing excellent model systems for studying the normal physiology and biochemistry of cells (e.g., metabolic studies, aging), the effects of drugs and toxic compounds on the cells, and mutagenesis and carcinogenesis. It is also used in drug screening and development, and large scale manufacturing of biological compounds (e.g., vaccines, therapeutic proteins). The major advantage of using cell culture for any of these applications is the consistency and reproducibility of results that can be obtained from using a batch of clonal cells.

- **Model systems**:

Cell cultures provide a good model system for studying 1) basic cell biology and biochemistry, 2) the interactions between disease-causing agents and cells, 3) the effects of drugs on cells, 4) the process and triggers for aging, and 5) nutritional studies.

- **Toxicity testing**:

Cultured cells are widely used alone or in conjunction with animal tests to study the effects of new drugs, cosmetics and chemicals

on survival and growth in a wide variety of cell types. Especially important are liver-and kidney-derived cell cultures.

- **Cancer Research**:

Since both normal cells and cancer cells can be grown in culture, the basic differences between them can be closely studied. In addition, it is possible, by the use of chemicals, viruses and radiation, to convert normal cultured cells to cancer causing cells. Thus, the mechanisms that cause the change can be studied. Cultured cancer cells also serve as a test system to determine suitable drugs and methods for selectively destroying types of cancer.

- **Virology**:

One of the earliest and major uses of cell culture is the replication of viruses in cell cultures (in place of animals) for use in vaccine production. Cell cultures are also widely used in the clinical detection and isolation of viruses, as well as basic research into how they grow and infect organisms.

- **Cell-Based Manufacturing**:

While cultured cells can be used to produce many important produces, three areas have generating the most interest. **First** is the large-scale production of viruses for use in vaccine production. These include vaccines for polio, rabies, chicken pox, hepatitis B and measles. **Second**, is the large scale production of cells that have been genetically engineered to produce proteins that have medicinal or commercial value. These include monoclonal antibodies, insulin, hormones, etc. **Third**, is the use of cells as replacement tissues and organs. Artificial skin for use in treating burns and ulcers is the first commercially available product. However, testing is underway on artificial organs such as pancreas, liver and kidney. A potential supply of replacement cells and tissues may come out of work currently being done with both embryonic and adult stem cells. These are cells that have the potential to

differentiate into a variety of different cells types. It is hoped that learning how to control the development of these cells may offer new treatment approaches for a wide variety of medical conditions.

- **Genetic counseling**:

Amniocentesis, a diagnostic technique that enables doctors to remove and culture fetal cells from pregnant women, has given doctors an important tool for the early diagnosis of fetal disorders. These cells can then be examined for abnormalities in their chromosomes and genes using karyotyping, chromosome painting and other molecular techniques.

- **Genetic Engineering**:

The ability to transfect or reprogram cultured cells with new genetic material (DNA and genes) has provided a major tool to molecular biologists wishing to study the cellular effects of the expression of theses genes (new proteins). These techniques can also be used to produce these new proteins in large quantity in cultured cells for further study. Insect cells are widely used as miniature cells factories to express substantial quantities of proteins that they manufacture after being infected with genetically engineered baculoviruses.

- **Gene Therapy**:

The ability to genetically engineer cells has also led to their use for gene therapy. Cells can be removed from a patient lacking a functional gene and the missing or damaged gene can then be replaced. The cells can be grown for a while in culture and then replaced into the patient. An alternative approach is to place the missing gene into a viral vector and then “infect” the patient with the virus in the hope that the missing gene will then be expressed in the patient’s cells.

- **Drug Screening and Development**:

Cell-based assays have become increasingly important for the pharmaceutical industry, not just for cytotoxicity testing but also for high

throughput screening of compounds that may have potential use as drugs. Originally, these cell culture tests were done in 96 well plates, but increasing use is now being made of 384 and 1536 well plates.

What are stem cells, and why are they important?

Stem cells have the remarkable potential to develop into many different cell types in the body during early life and growth. In addition, in many tissues they serve as a sort of internal repair system, dividing essentially without limit to replenish other cells as long as the person or animal is still alive. When a stem cell divides, each new cell has the potential either to remain a stem cell or become another type of cell with a more specialized function, such as a muscle cell, a red blood cell, or a brain cell. Stem cells are distinguished from other cell types by two important characteristics. First, they are unspecialized cells capable of renewing themselves through cell division, sometimes after long periods of inactivity. Second, under certain physiologic or experimental conditions, they can be induced to become tissue- or organ-specific cells with special functions. In some organs, such as the gut and bone marrow, stem cells regularly divide to repair and replace worn out or damaged tissues. In other organs, however, such as the pancreas and the heart, stem cells only divide under special conditions. Until recently, scientists primarily worked with two kinds of stem cells from animals and humans: embryonic stem cells and non-embryonic “somatic” or “adult” stem cells. The functions and characteristics of these cells will be explained in this document. Scientists discovered ways to derive embryonic stem cells from early mouse embryos nearly 30 years ago, in 1981. The detailed study of the biology of mouse stem cells led to the discovery, in 1998, of a method to derive stem cells from human embryos and grow the cells in the laboratory. These cells are called human embryonic stem cells. The embryos used in these studies were

created for reproductive purposes through *in vitro* fertilization procedures. When they were no longer needed for that purpose, they were donated for research with the informed consent of the donor. In 2006, researchers made another breakthrough by identifying conditions that would allow some specialized adult cells to be "reprogrammed" genetically to assume a stem cell-like state. This new type of stem cell, called induced pluripotent stem cells (IPSCs), will be discussed in a later section of this document.

Stem cells are important for living organisms for many reasons. In the 3- to 5-day-old embryo, called a blastocyst, the inner cells give rise to the entire body of the organism, including all of the many specialized cell types and organs such as the heart, lungs, skin, sperm, eggs and other tissues. In some adult tissues, such as bone marrow, muscle, and brain, discrete populations of adult stem cells generate replacements for cells that are lost through normal wear and tear, injury, or disease.

Given their unique regenerative abilities, stem cells offer new potentials for treating diseases such as diabetes and heart disease. However, much work remains to be done in the laboratory and the clinic to understand how to use these cells for cell-based therapies to treat disease, which is also referred to as regenerative or reparative medicine. Laboratory studies of stem cells enable scientists to learn about the cells' essential properties and what makes them different from specialized cell types. Scientists are already using stem cells in the laboratory to screen new drugs and to develop model systems to study normal growth and identify the causes of birth defects.

Research on stem cellscontinues to advance knowledge about how an organism develops from a single cell and how healthy cells replace damaged cells in adult organisms. Stem cell research is one of the most fascinating areas of contemporary biology, but, as with many

expanding fields of scientific inquiry, research on stem cells raises scientific questions as rapidly as it generates new discoveries.

What are the unique properties of all stem cells?

Stem cells differ from other types of cells in the body. All stem cells regardless of their source have three general properties: 1) they are capable of dividing and renewing themselves for long periods; 2) they are unspecialized; and 3) they can give rise to specialized cell types.

Stem cells are capable of dividing and renewing themselves for long periods. Unlike muscle cells, blood cells, or nerve cells which do not normally replicate themselves, stem cells may replicate many times, or proliferate. A starting population of stem cells that proliferates for many months in the laboratory can yield millions of cells. If the resulting cells continue to be unspecialized, like the parent stem cells, the cells are said to be capable of long-term self-renewal. Scientists are trying to understand two fundamental properties of stem cells that relate to their long-term self-renewal: Discovering the answers to these questions may make it possible to understand how cell proliferation is regulated during normal embryonic development or during the abnormal cell divisionthat leads to cancer. Such information would also enable scientists to grow embryonic and non-embryonic stem cells more efficiently in the laboratory. The specific factors and conditions that allow stem cells to remain unspecialized are of great interest to scientists. It has taken many years of trial and error to learn to derive and maintain stem cells in the laboratory without them spontaneously differentiating into specific cell types. For example, it took two decades to learn how to grow human embryonic stem cellsin the laboratory following the development of conditions for growing mouse stem cells. Likewise, scientists must first understand the signals that enable a non-embryonic (adult) stem cell population to proliferate and remain unspecialized before they will be

able to grow large numbers of unspecialized adult stem cells in the laboratory.

- **Stem cells are unspecialized:**

One of the fundamental properties of a stem cell is that it does not have any tissue-specific structures that allow it to perform specialized functions. For example, a stem cell cannot work with its neighbors to pump blood through the body (like a heart muscle cell), and it cannot carry oxygen molecules through the bloodstream (like a red blood cell). However, unspecialized stem cells can give rise to specialized cells, including heart muscle cells, blood cells, or nerve cells.

- **Stem cells can give rise to specialized cells:**

When unspecialized stem cells give rise to specialized cells, the process is called differentiation. While differentiating, the cell usually goes through several stages, becoming more specialized at each step. Scientists are just beginning to understand the signals inside and outside cells that trigger each step of the differentiation process. The internal signals are controlled by a cell's genes, which are interspersed across long strands of DNA and carry coded instructions for all cellular structures and functions. The external signals for cell differentiation include chemicals secreted by other cells, physical contact with neighboring cells, and certain molecules in the microenvironment. The interaction of signals during differentiation causes the cell's DNA to acquire epigenetic marks that restrict DNA expression in the cell and can be passed on through cell division. Many questions about stem cell differentiation remain. For example, are the internal and external signals for cell differentiation similar for all kinds of stem cells? Can specific sets of signals be identified that promote differentiation into specific cell types? Addressing these questions may lead scientists to find new ways to control stem cell differentiation in the laboratory, thereby growing

cells or tissues that can be used for specific purposes such as cell-based therapies or drug screening.

Adult stem cells typically generate the cell types of the tissue in which they reside. For example, a blood-forming adult stem cell in the bone marrow normally gives rise to the many types of blood cells. It is generally accepted that a blood-forming cell in the bone marrow which is called a hematopoietic stem cell cannot give rise to the cells of a very different tissue, such as nerve cells in the brain. Experiments over the last several years have purported to show that stem cells from one tissue may give rise to cell types of a completely different tissue. This remains an area of great debate within the research community. This controversy demonstrates the challenges of studying adult stem cells and suggests that additional research using adult stem cells is necessary to understand their full potential as future therapies.

- **Embryonic stem cells**

Embryonic stem cells, as their name suggests, are derived from embryos. Most embryonic stem cells are derived from embryos that develop from eggs that have been fertilized *in vitro* in an *in vitro* fertilizationclinic and then donated for research purposes with informed consent of the donors. They are *not* derived from eggs fertilized in a woman's body.

Embryonic stem cells grown in the laboratory

Growing cells in the laboratory is known as cell culture. Human embryonic stem cells (hESCs) are generated by transferring cells from a preimplantationstage embryo into a plastic laboratory culture dish that contains a nutrient broth known as culture medium. The cells divide and spread over the surface of the dish. The inner surface of the culture dish is typically coated with mouse embryonic skin cells that have been treated so they will not divide. This coating layer of cells is called a

feeder layer. The mouse cells in the bottom of the culture dish provide the cells a sticky surface to which they can attach. Also, the feeder cells release nutrients into the culture medium. Researchers have devised ways to grow embryonic stem cells without mouse feeder cells. This is a significant scientific advance because of the risk that viruses or other macromolecules in the mouse cells may be transmitted to the human cells.

The process of generating an embryonic stem cell line is somewhat inefficient, so lines are not produced each time cells from the preimplantation stage embryo are placed into a culture dish. However, if the plated cells survive, divide, and multiply enough to crowd the dish, they are removed gently and plated into several fresh culture dishes. The process of re-plating or subculturingthe cells is repeated many times and for many months. Each cycle of subculturing the cells is referred to as a passage. Once the cell line is established, the original cells yield millions of embryonic stem cells. Embryonic stem cells that have proliferated in cell culture for six or more months without differentiating, are pluripotent, and appear genetically normal are referred to as an embryonic stem cell line. At any stage in the process, batches of cells can be frozen and shipped to other laboratories for further culture and experimentation.

Embryonic stem cells stimulated to differentiate

As long as the embryonic stem cells in culture are grown under appropriate conditions, they can remain undifferentiated (unspecialized). But if cells are allowed to clump together to form embryoid bodies, they begin to differentiate spontaneously. They can form muscle cells, nerve cells, and many other cell types. Although spontaneous differentiation is a good indicator that a culture of embryonic stem cells is healthy, the process is uncontrolled and

therefore an inefficient strategy to produce cultures of specific cell types. So, to generate cultures of specific types of differentiated cells heart muscle cells, blood cells, or nerve cells, for example: scientists try to control the differentiation of embryonic stem cells. They change the chemical composition of the culture medium, alter the surface of the culture dish, or modify the cells by inserting specific genes. Through years of experimentation, scientists have established some basic protocols or "recipes" for the directed differentiationof embryonic stem cells into some specific cell types (Fig. 8).

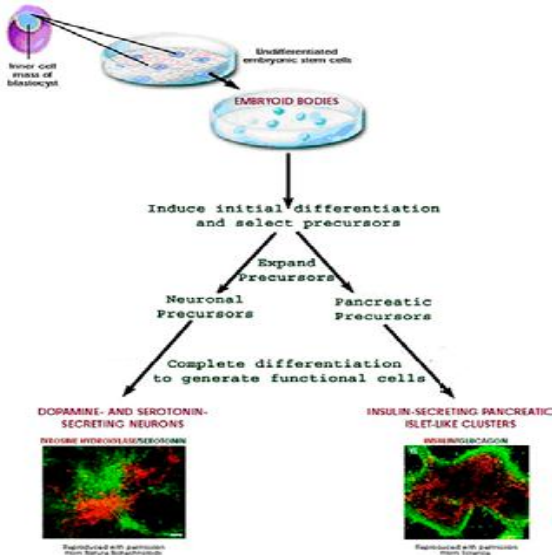


Figure 8: Directed differentiation of mouse embryonic stem cells

If scientists can reliably direct the differentiation of embryonic stem cells into specific cell types, they may be able to use the resulting, differentiated cells to treat certain diseases in the future. Diseases that might be treated by transplanting cells generated from human embryonic stem cells include diabetes, traumatic spinal cord injury,

Duchenne's muscular dystrophy, heart disease, and vision and hearing loss.

- **Adult stem cells:**

An adult stem cell is thought to be an undifferentiated cell, found among differentiated cells in a tissue or organ that can renew itself and can differentiate to yield some or all of the major specialized cell types of the tissue or organ. The primary roles of adult stem cells in a living organism are to maintain and repair the tissue in which they are found. Scientists also use the term somatic stem cell instead of adult stem cell, where somatic refers to cells of the body (not the germ cells, sperm or eggs). Unlike embryonic stem cells, which are defined by their origin (cells from the preimplantation-stage embryo), the origin of adult stem cells in some mature tissues is still under investigation. Research on adult stem cells has generated a great deal of excitement. Scientists have found adult stem cells in many more tissues than they once thought possible. This finding has led researchers and clinicians to ask whether adult stem cells could be used for transplants. In fact, adult hematopoietic, or blood-forming, stem cells from bone marrow have been used in transplants for 40 years. Scientists now have evidence that stem cells exist in the brain and the heart. If the differentiation of adult stem cells can be controlled in the laboratory, these cells may become the basis of transplantation-based therapies.

The history of research on adult stem cells began about 50 years ago. In the 1950s, researchers discovered that the bone marrow contains at least two kinds of stem cells. One population, called hematopoietic stem cells, forms all the types of blood cells in the body. A second population, called bone marrow stromal stem cells (also called mesenchymal stem cells, or skeletal stem cells by some) was discovered a few years later. These non-hematopoietic stem cells make up a small proportion of the stromal cell population in the bone marrow and can

generate bone, cartilage, and fat cells that support the formation of blood and fibrous connective tissue. In the 1960s, scientists who were studying rats discovered two regions of the brain that contained dividing cells that ultimately become nerve cells. Despite these reports, most scientists believed that the adult brain could not generate new nerve cells. It was not until the 1990s that scientists agreed that the adult brain does contain stem cells that are able to generate the brain's three major cell types - astrocytes and oligodendrocytes, which are non-neuronal cells, and neurons, or nerve cells.

- **Adult stem cell differentiation**:

As indicated above, scientists have reported that adult stem cells occur in many tissues and that they enter normal differentiation pathways to form the specialized cell types of the tissue in which they reside. Normal differentiation pathways of adult stem cells. In a living animal, adult stem cells are available to divide for a long period, when needed, and can give rise to mature cell types that have characteristic shapes and specialized structures and functions of a particular tissue. The following are examples of differentiation pathways of adult stem cells (Fig. 9) that have been demonstrated *in vitro*or *in vivo*.

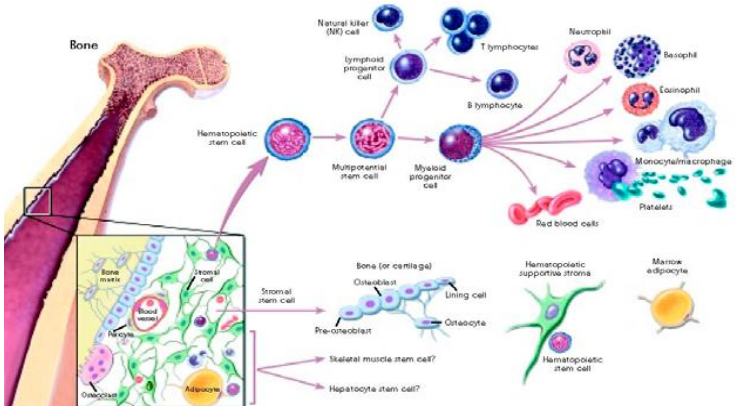


Figure 9: Hematopoietic and stromal stem cell differentiation

- **Hematopoietic stem** cells give rise to all the types of blood cells: red blood cells, B lymphocytes, T lymphocytes, natural killer cells, neutrophils, basophils, eosinophils, monocytes, and macrophages.
- **Mesenchymal stem cells** have been reported to be present in many tissues. Those from bone marrow (bone marrow stromal stem cells, skeletal stem cells) give rise to a variety of cell types: bone cells (osteoblasts and osteocytes), cartilage cells (chondrocytes), fat cells (adipocytes), and stromal cells that support blood formation. However, it is not yet clear how similar or dissimilar mesenchymal cells derived from non-bone marrow sources are to those from bone marrow stroma.
- **Neural stem cells** in the brain give rise to its three major cell types: nerve cells (neurons) and two categories of non-neuronal cells - **astrocytes** and **oligodendrocytes**.

- **Epithelial stem cells** in the lining of the digestive tract occur in deep crypts and give rise to several cell types: absorptive cells, goblet cells, Paneth cells, and enteroendocrine cells.
- **Skin stem cells** occur in the basal layer of the epidermis and at the base of hair follicles. The epidermal stem cells give rise to keratinocytes, which migrate to the surface of the skin and form a protective layer. The follicular stem cells can give rise to both the hair follicle and to the epidermis.

Transdifferentiation:

A number of experiments have reported that certain adult stem cell types can differentiate into cell types seen in organs or tissues other than those expected from the cells' predicted lineage (i.e., brain stem cells that differentiate into blood cells or blood-forming cells that differentiate into cardiac muscle cells, and so forth). Although isolated instances of transdifferentiation have been observed in some vertebrate species, whether this phenomenon actually occurs in humans is under debate by the scientific community. Instead of transdifferentiation, the observed instances may involve fusion of a donor cell with a recipient cell. Another possibility is that transplanted stem cells are secreting factors that encourage the recipient's own stem cells to begin the repair process. Even when transdifferentiation has been detected, only a very small percentage of cells undergo the process.

In a variation of transdifferentiation experiments, scientists have recently demon strated that certain adult cell types can be "reprogrammed" into other cell types *in vivo* using a well-controlled process of genetic modification. This strategy may offer a way to reprogram available cells into other cell types that have been lost or damaged due to disease. For example, one recent experiment shows how pancreatic beta cells, the insulin-producing cells that are lost or

damaged in diabetes, could possibly be created by reprogramming other pancreatic cells. By “re-starting” expression of three critical beta cell genes in differentiated adult pancreatic exocrine cells, researchers were able to create beta cell-like cells that can secrete insulin. The reprogrammed cells were similar to beta cells in appearance, size, and shape; expressed genes characteristic of beta cells; and were able to partially restore blood sugar regulation in mice whose own beta cells had been chemically destroyed. While not transdifferentiation by definition, this method for reprogramming adult cells may be used as a model for directly reprogramming other adult cell types.

In addition to reprogramming cells to become a specific cell type, it is now possible to reprogram adult somatic cells to become like embryonic stem cells (**induced pluripotent stem cells, iPSCs**) through the introduction of embryonic genes. Thus, a source of cells can be generated that are specific to the donor, thereby avoiding issues of histocompatibility, if such cells were to be used for tissue regeneration. However, like embryonic stem cells, determination of the methods by which iPSCs can be completely and reproducibly committed to appropriate cell lineages is still under investigation.

Similarities and Differences between Embryonic and Adult stem cells

Human embryonic and adult stem cells each have advantages and disadvantages regarding potential use for cell-based regenerative therapies. One major difference between adult and embryonic stem cells is their different abilities in the number and type of differentiated cell types they can become. Embryonic stem cells can become all cell types of the body because they are pluripotent. Adult stem cells are thought to be limited to differentiating into different cell types of their tissue of origin. Embryonic stem cells can be grown relatively easily in culture. Adult stem cells are rare in mature tissues, so isolating these cells from

an adult tissue is challenging, and methods to expand their numbers in cell culture have not yet been worked out. This is an important distinction, as large numbers of cells are needed for stem cell replacement therapies. Scientists believe that tissues derived from embryonic and adult stem cells may differ in the likelihood of being rejected after transplantation. We don't yet know for certain whether tissues derived from embryonic stem cells would cause transplant rejection, since relatively few clinical trials have tested the safety of transplanted cells derived from hESCS. Adult stem cells, and tissues derived from them, are currently believed less likely to initiate rejection after transplantation. This is because a patient's own cells could be expanded in culture, coaxed into assuming a specific cell type (differentiation), and then reintroduced into the patient. The use of adult stem cells and tissues derived from the patient's own adult stem cells would mean that the cells are less likely to be rejected by the immune system. This represents a significant advantage, as immune rejection can be circumvented only by continuous administration of immunosuppressive drugs, and the drugs themselves may cause deleterious side effects.

Induced pluripotent stem cells (iPSCs) are adult cells that have been genetically reprogrammed to an embryonic stem cell–like state by being forced to express genes and factors important for maintaining the defining properties of embryonic stem cells. Although these cells meet the defining criteria for pluripotent stem cells, it is not known if iPSCs and embryonic stem cells differ in clinically significant ways. Mouse iPSCs were first reported in 2006, and human iPSCs were first reported in late 2007. Mouse iPSCs demonstrate important characteristics of pluripotent stem cells, including expressing stem cell markers, forming tumors containing cells from all three germ layers, and being able to contribute too many different tissues when injected into mouse embryos at a very

early stage in development. Human iPSCs also express stem cell markers and are capable of generating cells characteristic of all three germ layers. Although additional research is needed, iPSCs are already useful tools for drug development and modeling of diseases, and scientists hope to use them in transplantation medicine. Viruses are currently used to introduce the reprogramming factors into adult cells, and this process must be carefully controlled and tested before the technique can lead to useful treatments for humans. In animal studies, the virus used to introduce the stem cell factors sometimes causes cancers. Researchers are currently investigating non-viral delivery strategies. In any case, this breakthrough discovery has created a powerful new way to "de-differentiate" cells whose developmental fates had been previously assumed to be determined. In addition, tissues derived from iPSCs will be a nearly identical match to the cell donor and thus probably avoid rejection by the immune system. The iPSC strategy creates pluripotent stem cells that, together with studies of other types of pluripotent stem cells, will help researchers learn how to reprogram cells to repair damaged tissues in the human body.

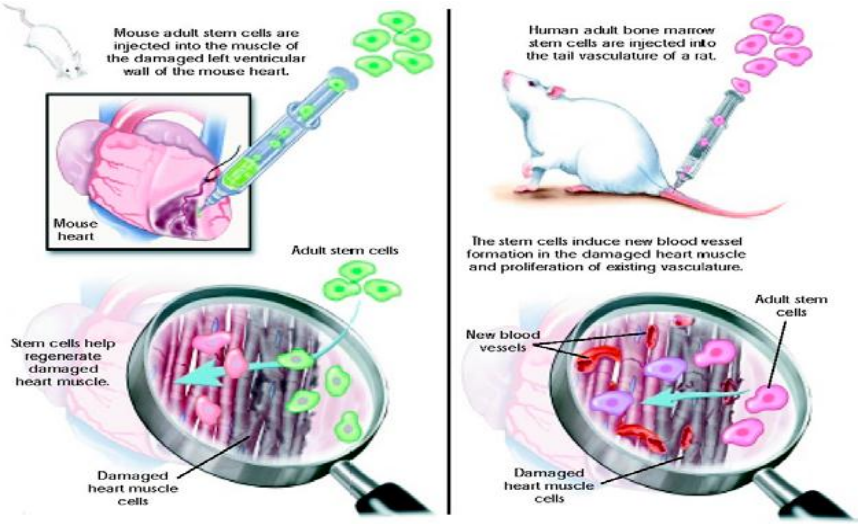
Potential uses of human stem cells and the obstacles

There are many ways in which human stem cells can be used in research and the clinic. Studies of human embryonic stem cellswill yield information about the complex events that occur during human development. A primary goal of this work is to identify how undifferentiatedstem cells become the differentiated cells that form the tissues and organs. Scientists know that turning geneson and off are central to this process. Some of the most serious medical conditions, such as cancer and birth defects, are due to abnormal cell division and differentiation. A more complete understanding of the genetic and molecular controls of these processes may yield information about how

such diseases arise and suggest new strategies for therapy. Predictably controlling cell proliferation and differentiation requires additional basic research on the molecular and genetic signals that regulate cell division and specialization. While recent developments with iPS cells suggest some of the specific factors that may be involved, techniques must be devised to introduce these factors safely into the cells and control the processes that are induced by these factors.

Human stem cells are currently being used to test new drugs. New medications are tested for safety on differentiated cells generated from human pluripotent cell lines. Other kinds of cell lines have a long history of being used in this way. Cancer cell lines, for example, are used to screen potential anti-tumor drugs. The availability of pluripotent stem cells would allow drug testing in a wider range of cell types. However, to screen drugs effectively, the conditions must be identical when comparing different drugs. Therefore, scientists will have to be able to precisely control the differentiation of stem cells into the specific cell type on which drugs will be tested. Current knowledge of the signals controlling differentiation falls short of being able to mimic these conditions precisely to generate pure populations of differentiated cells for each drug being tested. Perhaps the most important potential application of human stem cells is the generation of cells and tissues that could be used for cell-based therapies. Today, donated organs and tissues are often used to replace ailing or destroyed tissue, but the need for transplantable tissues and organs far outweighs the available supply. Stem cells, directed to differentiate into specific cell types, offer the possibility of a renewable source of replacement cells and tissues to treat diseases including Alzheimer's disease, spinal cord injury, stroke, burns, heart disease, diabetes, osteoarthritis, and rheumatoid arthritis. For example, it may become possible to generate healthy heart muscle cells in the laboratory and then transplant those cells into patients with

chronic heart disease. Preliminary research in mice and other animals indicates that bone marrow stromal cells, transplanted into a damaged heart, can have beneficial effects. Whether these cells can generate heart muscle cells or stimulate the growth of new blood vessels that repopulate the heart tissue, or help via some other mechanism is actively under investigation. For example, injected cells may repair by secreting growth factors, rather than actually incorporating into the heart. Promising results from animal studies have served as the basis for a small number of exploratory studies in humans. Other recent studies in cell culturesystems indicate that it may be possible to direct the differentiationof embryonic stem cells or adult bone marrow cells into heart muscle cells (Fig. 10).


Mouse adult stem cells are injected into the muscle of the damaged left ventricular wall of the mouse heart.
Mouse heart
Adult stem cells
Stem cells help regenerate damaged heart muscle.
Damaged heart muscle cells
Human adult bone marrow stem cells are injected into the tail vasculature of a rat.
The stem cells induce new blood vessel formation in the damaged heart muscle and proliferation of existing vasculature.
Adult stem cells
New blood vessels
Damaged heart muscle cells

.**Figure 10: Strategies to repair heart muscle with adult stem cells**

To realize the promise of novel cell-based therapies for such pervasive and debilitating diseases, scientists must be able to manipulate stem cells so that they possess the necessary characteristics for

successful differentiation, transplantation, and engraftment. The following is a list of steps in successful cell-based treatments that scientists will have to learn to control to bring such treatments to the clinic. To be useful for transplant purposes, stem cells must be reproducibly made to:

- Proliferate extensively and generate sufficient quantities of cells for making tissue.
- Differentiate into the desired cell type(s).
- Survive in the recipient after transplant.
- Integrate into the surrounding tissue after transplant.
- Function appropriately for the duration of the recipient's life.
- Avoid harming the recipient in any way.
- Also, to avoid the problem of immune rejection, scientists are experimenting with different research strategies to generate tissues that will not be rejected.

To summarize, stem cells offer exciting promise for future therapies, but significant technical hurdles remain that will only be overcome through years of intensive research.

Genetic manipulation of animals (animal models in research)

A transgenic animal is an animal whose hereditary DNA has been augmented by addition of DNA from a source other than parental germplasm through recombinant DNA techniques. Transfer of genes or gene constructs allows for the manipulation of individual genes rather than entire genoms. There have been dramatic advances in gene transfer technology in the last two decades since the first successful transfer was carried out in mice in 1980 (Palmiter*et al*., 1982; Jaenisch, 1988). The technique has now become routine in the mouse and resulting transgenic mice are able to transmit their transgenic to their offspring thereby allowing a large number of transgenic animals to be produced.

Successful production of transgenic livestock has been reported for fish, pigs, sheep, rabbits and cattle. The majority of gene transfer studies in livestock have, however, been carried out in the pig. Although transgenic cattle and sheep have been successfully produced, the procedure is still inefficient in these species (Niema*et al*., 1994). Why are these animals being produced? How can man benefit from such modifications?

To know some of the common reasons:

For Medical Purpose:

Normal physiology and development:

Transgenic animals can be specifically designed to allow the study of how genes are regulated, and how they affect the normal functions of the body and its development, e.g., study of complex factors involved in growth such as insulin-like growth factor. By introducing genes from other species that alter the formation of this factor and studying the biological effects that result, information is obtained about the biological role of the factor in the body.

- **Study of disease:** Many transgenic animals are designed to increase our understanding of how genes contribute to the development of disease. These are specially made to serve as models for human diseases so that investigation of new treatments for diseases is made possible. Today transgenic models exist for many human diseases such as cancer, cystic fibrosis, rheumatoid arthritis and Alzheimer's.
- **Biological products:** Medicines required to treat certain human diseases can contain biological products, but such products are often expensive to make. Transgenic animals that produce useful biological products can be created by the introduction of the portion of DNA (or genes) which codes for a particular product such as human protein (**α**-1-antitrypsin) used to treat emphysema. Similar attempts are being made for treatment of phenylketonuria (PKU) and cystic

fibrosis. In 1997, the first transgenic cow, Rosie, produced human protein-enriched milk (2.4 grams per litre). The milk contained the human alpha-lactalbumin and was nutritionally a more balanced product for human babies than natural cow-milk.

- **Vaccine safety:** Transgenic mice are being developed for use in testing the safety of vaccines before they are used on humans. Transgenic mice are being used to test the safety of the polio vaccine. If successful and found to be reliable, they could replace the use of monkeys to test the safety of batches of the vaccine.
- **Chemical safety testing:** This is known as toxicity/safety testing. The procedure is the same as that used for testing toxicity of drugs. Transgenic animals are made that carry genes which make them more sensitive to toxic substances than non-transgenic animals. They are then exposed to the toxic substances and the effects studied. Toxicity testing in such animals will allow us to obtain results in less time.

Animal Breeding and Transgenic Animals

Transgenesis offers considerable opportunity for advances agriculture. In livestock, the ability to insert new genes for such economically important characteristics as fecundity, resistance to or tolerance of other environmental stresses would represent a major breakthrough in the breeding of commercially superior stock. Another opportunity that transgenic technology could provide is in the production of clotting factors in the milk of domestic livestock. The genes coding for these proteins have been identified and the human factor IX construct has been successfully introduced into sheep and expression achieved in sheep milk (Clark *et al*., 1990). Moreover, the founder animal has been shown to be able to transmit the trait to its offspring (Niemane*t al*., 1994). To date, the majority of genes transferred into

sheep have been growth hormone encoding gene constructs. Unfortunately, in most cases the elevated growth hormone levels have resulted into a clinical diabetes situation leading to an early death of the transgenic sheep (Rexroad*et al*., 1990). The first reports of the production of transgenic animals created a lot of excitement among biological scientists. In the field of animal breeding, there were diverse opinions on how the technology might affect livestock genetic improvement programmes. Some (Ward *et al*., 1982) believed that it would result in total re-organization of conventional animal breeding theory while others (Schuman and Shoffner, 1982) considered the technology as an extension of current animal breeding procedures which, by broadening the gene pool, would make new and novel genotypes available for selection. Application of the technology in animal improvement is still far from being achieved. However, consideration needs to be given to its potential role in this field. Smith *et al* (1987) presented a comprehensive evaluation of strategies for developing, testing, breeding and disseminating transgenic livestock in the context of quantitative improvement of economic traits. An important contribution of transgenic technology is in the area of basic research to study the role of genes in the control of physiological processes. The understanding of the molecular control of life processes has important implications for both medicine and agriculture. For example, the generation (through mutation of an endogenous gene) of an organism which lacks a specific gene is a powerful tool to investigate the function of the gene product. This type of genetic analysis has been facilitated by the availability of *in vitro* cultures of embryonic stem cells from mice (Bradley *et al*., 1984).

Recent advances in *in vitro* technology (*in vitro* fertilization and maturation) will increase the number of zygotes available for gene transfer purposes. This, plus the utilization of embryonic stem cell

(Stice*et al*., 1994) and primodial germ cell technologies should enhance the efficiency of gene transfer in cattle and sheep considerably.

Methods of genetic manipulation in animals:

A transgenic animal is one that carries a foreign gene that has been deliberately inserted into its genome. The foreign gene is constructed using recombinant DNA methodology. Two methods of producing transgenic animals are widely used: (1) transforming embryonic stem cells (ES cells) growing in tissue culture with the desired DNA and (2) injecting the desired gene into the pronucleus of a fertilized egg. (3) Desirab Retrovirus-mediated transgenesis. (4) Pronuclear microinjection (Fig. 11). (5) Sperm-mediated transfer.

Genes from one species are transferred to other animals or species to improve the productivity of livestock. Faster growth rates, leaner growth patterns, more resistance to disease, increased milk production, more efficient metabolism, and transferring antimicrobial genes to farm animals are some of the goals of transgenic animal researchers. These include laboratory culture of large numbers of viable embryos for nonsurgical transfer to surrogate mothers, development of methods for sexing sperm and embryos, cloning embryos by nuclear transplantation and gene transfer to create livestock with superior performance traits. In all cases material progress will depend upon a deeper understanding of the underlying physiological and developmental control mechanisms and public confidence that due regard is being paid to animal welfare, and to social and environmental implications. Genetic improvement of livestock depends on access to genetic variation and effective methods for exploiting this variation. Genetic diversity constitutes a buffer against changes in the environment and is a key in selection and breeding for adaptability and production on a range of environments.

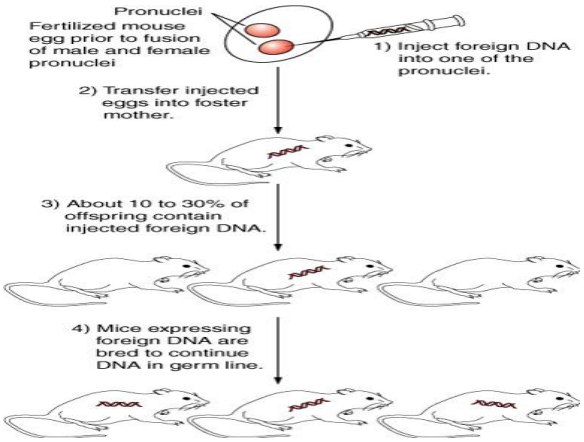

Pronuclei
Fertilized mouse egg prior to fusion of male and female pronuclei
1) Inject foreign DNA into one of the pronuclei.
2) Transfer injected eggs into foster mother.
3) About 10 to 30% of offspring contain injected foreign DNA.
4) Mice expressing foreign DNA are bred to continue DNA in germ line.

Figure 11: Pronuclear microinjection

In developed countries, breeding programmes are based upon performance recording and this has led to substantial improvements in animal production. Developing countries have distinct disadvantages for setting up successful breeding programmes: infrastructure needed for performance testing is normally lacking because herd sizes are normally small and variability between farms, farming systems and seasons are large; reproductive efficiency is low, due mainly to poor nutrition, especially in cattle; and communal grazing precludes implementation of systematic breeding and animal health programmes.

A brief history of cloning

- **1938 – First idea of cloning:** Hans Spemann proposes a "fantastic experiment" to replace the nucleus of an egg cell with the nucleus of another cell and to grow an embryo from such an egg;
- **1952 – An attempt to clone a *Ranapipiens* frog:** Robert Briggs and Thomas King; the scientists collect the nucleus from a frog egg cell with a pipette and replace it with the nucleus taken from a cell of a frog embryo; the experiment is not successful;

- **1970 – *Xenopuslaevis* frog:** John B. Gurdon is successful: he clones a frog, but its development only reaches the stage of tadpole. Despite attempts, he never manages to obtain an adult specimen. For many years, his achievement is questioned, especially in light of unsuccessful attempts to clone mammals;
- **1981 – Karl Illmenese and Peter Hope clone a mouse:** They take the nucleus not from an adult specimen, but from a mouse embryo;
- **1994 – Neal First tries to clone a sheep:** He takes the nucleus from an embryonic cell. He obtains a sheep embryo that develops 120 cells;
- **1995 – Two sheep are cloned (Moran and Megan):** These had been the first animals cloned from differentiated cells obtained by means of a pioneering method of nuclei transfer. However, the cells from which the nucleus was taken did not come directly from another living animal, but from a cell culture. The ones who achieved that were Ian Wilmut and Keith Campbell;
- **1996 – The first mammal cloned from a cell taken from an adult animal** – Dolly the sheep. Creators: Ian Wilmut and Keith Campbell;
- **1998 – The first cloned mouse** (it was called Cumulina);
- **2000 – The first cloned rhesus monkey**;
- **2000 – The first cloned pig** (or even five pigs);
- **2001 – A buffalo and a cow cloned**;
- **2001 – A cat cloned** (it was called Copy Cat);
- **2002** – KonradHochedlinger and Rudolf Jaenisch clone **mice from T lymphocytes**;
- **2003 – A rabbit is cloned** in France and Southern Korea;
- **2003 – A mule is cloned**. It was achieved by the companies Idaho Gem and Utah Pioneer;
- **2003 – A deer** (Dewey), a horse (Prometea) and a rat (Ralph) cloned;
- **2004 – Fruit flies cloned**;

- **2005 – An Afghan hound** (Snuppy) cloned;
- **2007 – A wolf cloned**; South Korean scientists obtained two female wolves (Snuwolf and Snuwolffy);
- **2008 – A Labrador dog cloned**;
- **2009 – The first animal from an extinct species cloned**: Pyrenean ibex. The animal lived seven minutes. It died of lung malformations;
- **2009 – A camel female cloned (Injaz)**; Injaz was created from ovarian cells of a female killed for meat in 2005.
- **2009 - Samrupa, the world's first** Murrah buffalo **calf cloned** using a simple "Hand guided cloning technique" was born in February, 2009 at National Dairy Research Institute (NDRI), Karnal, India, but died due to a lung infection five days after she was born.
- **2009– Garima-I, a buffalo calf cloned** using an "Advanced Hand guided Cloning Technique" was born in June, 2009 at the NDRI. Two years later in 2011, she died of a heart failure.
- **2009 - Garima-II, another cloned** calf was born in August, 2010. **A cloned male buffalo** calf Shresth was born in August, 2010 at the NRDI.
- **2010 - Got, the first Spanish Fighting Bull was cloned** by Spanish scientists.
- **2013 - This buffalo** was inseminated with frozen-thawed semen of a progeny tested bull and gave birth to a female calf, Mahima in January, 2013.

Cloning

The procedure for obtaining organisms with the same genetic information. You need to collect an egg cell from a donor (a female sheep, mouse or cat). Then you need to carefully remove the nucleus from the cell and collect another cell from the skin, udder or other tissue from another male or female donor of the same species (i.e. from

another sheep, mouse or cat). From this cell, you also need to remove the cell nucleus and place it in the empty egg. The egg obtained in such a way needs to be treated with a gentle electric shock. The egg should begin to divide and grow into a multicellular embryo. At this stage, the embryo needs to be implanted into the uterus of a surrogate mother. If the pregnancy develops and the animal is born (Fig. 12).

(c

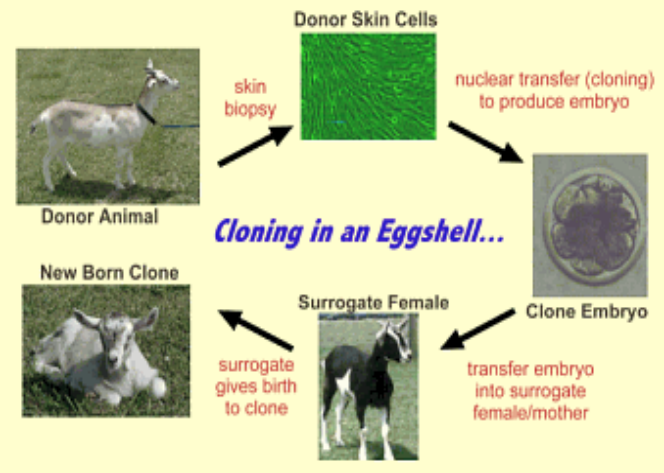


Figure 12: Steps in cloning simplified

Creating a clone of your favorite animal seems like a great way to insure your pet will be with you forever. Although this might be a goal of cloning, it is not the primary focus of biotech specialists. Commercialization of cloning allows desirable traits to be reliably propagated. Animal breeders are able to clone animals with superior traits such as cows with high milk production or champion racehorses. Embryo twinning (splitting embryos in half) was the first method of cloning used to produce identical twin cattle. Since the twins are the result of mixing the genetic material from two parents, the exact genetic make-up of the animal is not known until it has matured. Dolly (the very

famous sheep that was the first mammal ever cloned in the lab), however, was created from a single cell, not an embryo. DNA from a donor cell is inserted into an egg that has had its own DNA removed. It is a very delicate and difficult process. So far, animals successfully cloned include sheep, goats, pigs, cattle, cats, deer and dogs.

One can imagine future uses of cloning that could include using preserved DNA to help maintain endangered species or even recover extinct species!

- **Limits to Cloning**: The donor cell must come from a living organism: an organism is also shaped by its environment, success rate for cloning is very low and clones may be old before their time.
- **The future of cloning**: preservation of endangered animals, studying the effect of drugs etc on duplicates, improve agricultural production (Fig. 13)

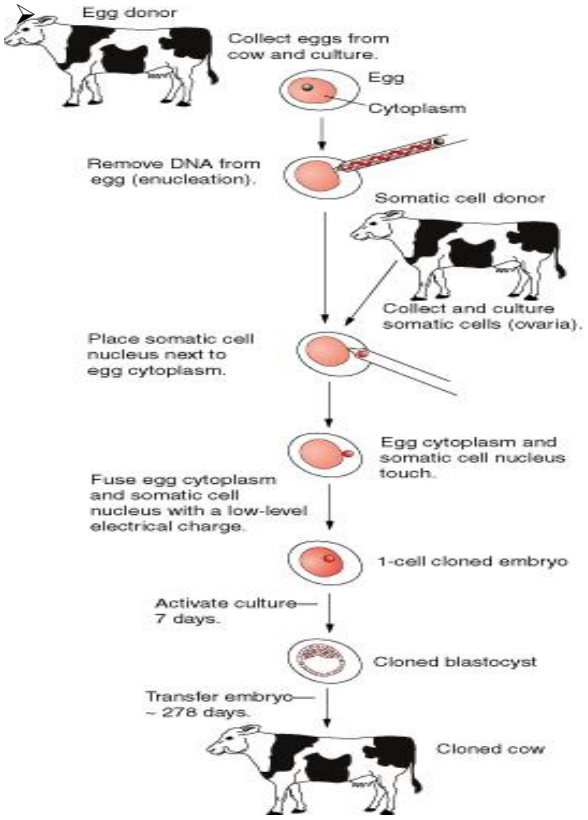

Egg donor
Collect eggs from cow and culture.
Egg
Cytoplasm
Remove DNA from egg (enucleation).
Somatic cell donor
Collect and culture somatic cells (ovaria).
Place somatic cell nucleus next to egg cytoplasm.
Egg cytoplasm and somatic cell nucleus touch.
Fuse egg cytoplasm and somatic cell nucleus with a low-level electrical charge.
1-cell cloned embryo
Activate culture— 7 days.
Cloned blastocyst
Transfer embryo— ~ 278 days.
Cloned cow

Figure 13: Cloning for improved agricultural productionApplications of transgenic animals.

Use of Animals in Research

- Animals play a vital role in primary research. The use of animal models permits more rapid assessment of the effects of new medical treatments and other products. Computer models and *in vitro* studies of cell cultures are often used as supplements to animal research, but they can't entirely duplicate the results in living organisms. Recent developments in animal biotechnology

have changed medicine, agriculture, and the efforts to preserve endangered animals.

- For a new product to be approved for human use, the manufacturer must first demonstrate that it is safe for use. Trials are required on cell cultures, in live animals, and on human subjects. Testing on live animal models requires that two or more species be used because different effects are observed in different animals. If problems are detected in the animal tests, human subjects are never recruited for trials. The animals used most often are pure-bred mice and rats, but other species are also used. Another extremely valuable research animal is the zebrafish, a hardy aquarium fish. Dogs are used for the study of cancer, heart disease and lung disorders. HIV and AIDS research is conducted on monkeys and chimpanzees.
- Animal research is very heavily regulated. The Animal Welfare Act sets standards concerning the housing, feeding, cleanliness and medical care of research animals. Veterinarians also conduct research which has led to new cancer treatments for pets and studies in their adaptations for humans.
- Animal Models
- Mice
- Rats
- Zebrafish (3 month generation time, 200 progeny, complete embryogenesis in 120 hrs)
- Dogs (lungs and cardiovascular system)
- Cats
- Pigs (PPL Therapeutics- delete a gene which causes hyperacute rejection of pig-to-human organ transplantation)
- Primates (HIV and AIDs research, geriatric research)
- Bioengineering Mosquitoes to Prevent Malaria

- Cloned in a gene that prevents the parasite from traversing the midgut; blocking the continuation of its life cycle
- Developed an antibody that prevents the parasite from entering the mosquito's salivary gland

Pharming: Pharming is not just a misspelled word ! The term "pharming" comes from a combination of the words "farming" and "pharmaceuticals" - a blending of the basic methods of agriculture with advanced biotechnology.

Gene pharming is a technology that scientists use to alter an animal's own DNA, or to splice in new DNA, called a transgene, from another species. In pharming, these genetically modified (transgenic) animals are mostly used to make human proteins that have medicinal value. The protein encoded by the transgene is secreted into the animal's milk, eggs or blood, and then collected and purified (Fig. 14). One interesting GMO organism that has been in the news lately is the "glowing fish." GloFish ™ fluorescent zebra fish were specially bred to help detect environmental pollutants.

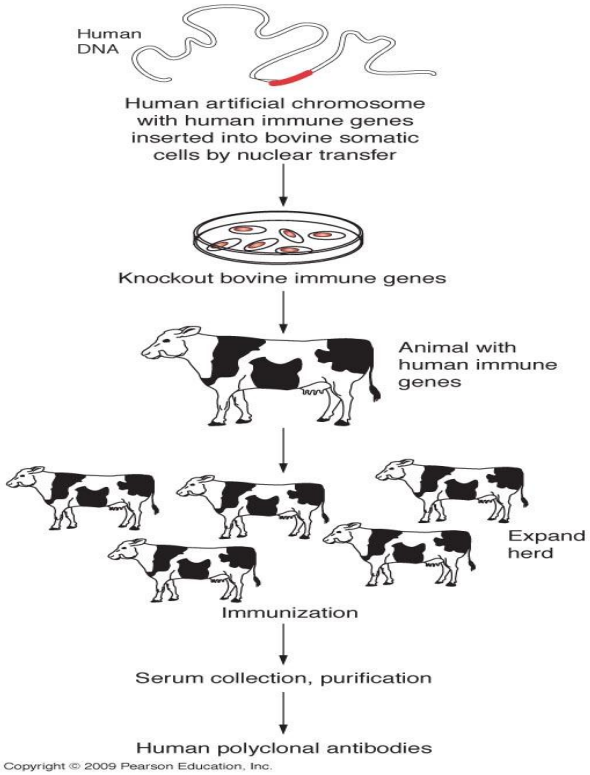

Human
DNA
Human artificial chromosome with human immune genes inserted into bovine somatic cells by nuclear transfer
Knockout bovine immune genes
Animal with human immune genes
Expand herd
Immunization
Serum collection, purification
Human polyclonal antibodies
Copyright © 2009 Pearson Education, Inc.

Figure 14: Pharming-Animal used to produce human polyclonal antibodies

Improving Agricultural Products with Transgenics

- Faster growth rates or leaner growth patterns (improve the product), more product
- Increase nutritional content-lactoferrin
- Turning the animals into efficient grazers
- Transfer antimicrobial genes to farm animals

Transgenic Animals as Bioreactors

- Biosteel otherwise known as spider silk, cloned into goat milk ("silkmilk" goats) (Fig. 15).
- -Goats reproduce faster than cows and are cheaper than cows.
- -Hens also make good bioreactors in that they are cheap and a lot of eggs are produced at one time.

Figure 15: Goats act as bioreactors

Genetically Engineered Insulin:

Management of adult-onset diabetes is possible by taking insulin at regular time intervals. What would a diabetic patient do if enough human-insulin was not available? If you discuss this, you would soon realize that one would have to isolate and use insulin from other animals. Would the insulin isolated from other animals be just as effective as that secreted by the human body itself and would it not elicit an immune response in the human body? Now, imagine if bacterium were available that could make human insulin. Suddenly the whole process becomes so simple. You can easily grow a large quantity of the bacteria and make as much insulin as you need. Think about whether insulin can be orally administered to diabetic people or not. Why? Insulin used for diabetes was earlier extracted from pancreas of slaughtered cattle and pigs. Insulin from an animal source though caused some

patients to develop allergy or other types of reactions to the foreign protein. Insulin consists of two short polypeptide chains: chain A and chain B, which are linked together by disulphide bridges (Fig. 16). In mammals, including humans, insulin is synthesized as a pro-hormone (like a pro-enzyme, the pro-hormone also needs to be processed before it becomes a fully mature and functional hormone) which contains an extra stretch called the C peptide. This C peptide is not present in the mature insulin and is removed during maturation into insulin.The main challenge for production of insulin using rDNA techniques was getting insulin assembled into a mature form. In 1983, Eli Lilly an American company prepared two DNA sequences corresponding to A and B, chains of human insulin and introduced them in plasmids of E. coli to produce insulin chains. Chains A and B were produced separately, extracted and combined by creating disulfide bonds to form human insulin.

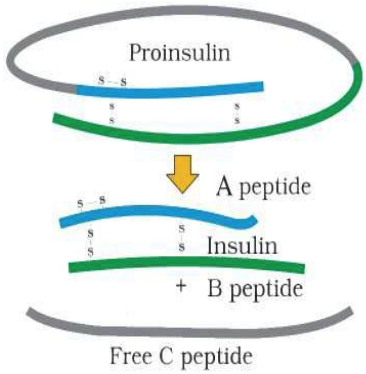


Figure 16: Genetically engineered insulin

Transgenic and cloning (Translational Significance):

The reality itsvery expensive technology. Technologies still need to be refined large numbers of repetitions required to produce viable

offspring in animals. Applications currently very limited predominantly used for biomedical purposes.

Gene Therapy:

If a person is born with a hereditary disease, can a corrective therapy be taken for such a disease? Gene therapy is an attempt to do this. Gene therapy is a collection of methods that allows correction of a gene defect that has been diagnosed in a child/embryo. Here genes are inserted into a person's cells and tissues to treat a disease. Correction of a genetic defect involves delivery of a normal gene into the individual or embryo to take over the function of and compensate for the non-functional gene. In simple, Genes are inserted into cells and tissues of an individual to correct certain hereditary diseases.It involves the delivery of a normal gene into the individual or embryo to replace the defective mutant allele of the gene.Viruses which attack the host and introduce their genetic material into host are all used as vectors.The first clinical gene therapy was given in 1990 to a four year old girl with adenosine deaminase (ADA) deficiency. ADA deficiency can be cured by bone marrow transplantation in some children but it is not completely curative. For gene therapy, lymphocytes were grown in a cultural and functional ADA. cDNA is then introduced into these lymphocytes. These lymphocytes are then transferred into the body of the patient; the patient requires periodically infusion of such genetically engineered lymphocytes. If a functional gene is introduced into the bone marrow cells at early embryonic stage, it would be permanent cure.

Monoclonal Antibodies:

Production of Monoclonal antibodies (Mabs) (Fig. 17). Used to treat cancer, heart disease, and transplant rejection. HUMANIZED monoclonal antibodies were developed to prevent the human anti-mouse antibody (HAMA) response

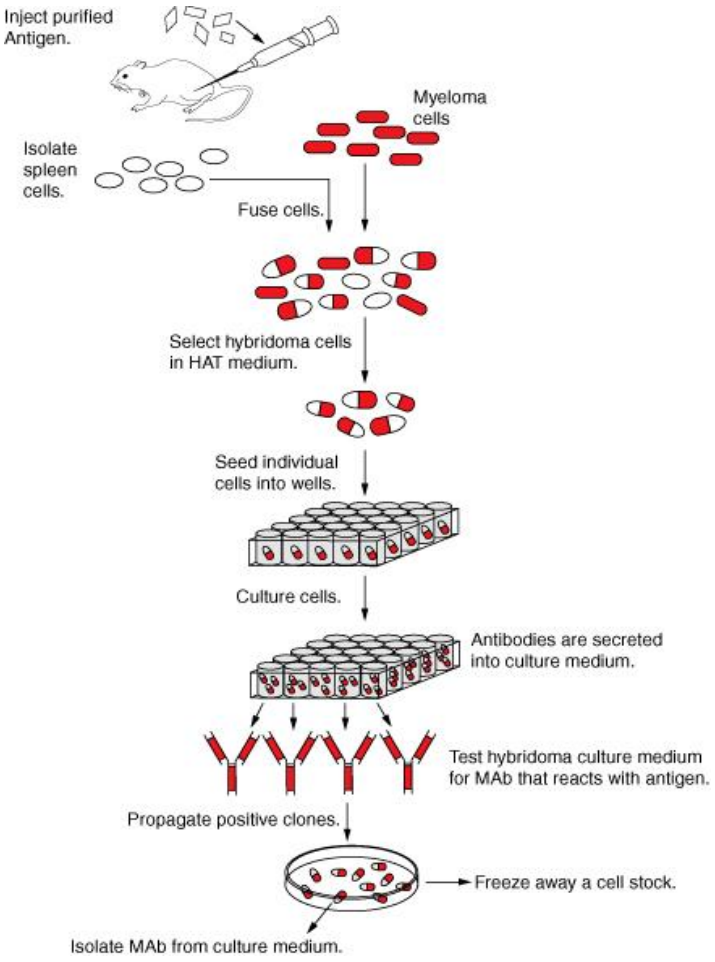

Inject purified Antigen.
Myeloma cells
Isolate spleen cells.
Fuse cells.
Select hybridoma cells in HAT medium.
Seed individual cells into wells.
Culture cells.
Antibodies are secreted into culture medium.
Test hybridoma culture medium for MAb that reacts with antigen.
Propagate positive clones.
Freeze away a cell stock.
Isolate MAb from culture medium.

Figure 17: Production of monoclonal antiboides

Economic impact in developing countries:

The developing world is grossly unprepared for the new technological and economic opportunities, challenges and risks that lie on the horizon. Although it is hoped that biotechnology will improve the

life of every person in the world and allow more sustainable living, crucial decisions may be dictated by commercial considerations and the socioeconomic goals that society considers to be the most important.

The use of biotechnology will lead to a distinct shift in the economic returns from livestock. Though the role of livestock in ensuring nutritional security is recognized in mixed crop-livestock systems, the importance of livestock goes beyond direct food production. Livestock supply draught power and organic manure to the crop sector, and hides, skins, bones, blood and fiber are used in many industries. Thus, livestock are an important source of income and employment, helping to alleviate poverty and smooth the income distribution among small landholders and the landless, who constitute the bulk of the rural population and the majority of livestock owners. In addition, livestock can easily be converted into cash and thus act as a cushion against crop failure, particularly in less favored environments. By enabling crop residues and by-products to be used as fodder, livestock production contributes positively to the environment.In developed countries livestock accounts for more than half of agricultural production, while in developing countries the share is about one-third. This latter share, however, is rising quickly because of rapid increases in livestock production resulting from population growth, urbanization, changes in lifestyles and dietary habits and increasing disposable incomes.

In most developing countries, biotechnological applications relating to livestock need to be suitable for animal owners who are resource-poor small-scale operators who own little or no land and few animals. Using technology to support livestock production is an integral part of viable agriculture in multi-enterprise systems. Livestock are part of a fragile ecosystem and a rich source of animal biodiversity, as local species and breeds possess genes and traits of excellence. Molecular markers are increasingly being used to identify and select the particular

genes that lead to these desirable traits and it is now possible to select superior germ plasm and disseminate it using artificial insemination, embryo transfer and other assisted reproductive technologies. These technologies have been used in the genetic improvement of livestock, particularly in cattle and buffaloes, and the economic returns are significant. However, morbidity and mortality among animals produced using assisted reproductive technologies lead to high economic losses, so the principal application of animal biotechnology at present is in the production of cheap and dependable diagnostic kits and vaccines. Several obstacles limit the application of biotechnology at present: there is a lack of infrastructure and insufficient manpower, so funding is needed if resource-poor farmers are to benefit from biotechnology.

Ethical Issues; (biosafety and regulatory issues):

Biosafety and Ethics

All procedures involved in the collection of human material for culture must be passed by the relevant hospital ethics committee. A form will be required for the patient to sign authorizing research use of the tissue, and preferably disclaiming any ownership of any materials derived from the tissue [Freshney, 2002, 2005]. The form should have a brief layman's description of the objectives of the work and the name of the lead scientist on the project. The donor should be provided with a copy. All human material should be regarded as potentially infected and treated with caution. Samples should be transported securely in double-wrapped waterproof containers; they and derived cultures should be handled in a Class II biosafety cabinet and all discarded material autoclaved, incinerated, or chemically disinfected. Each laboratory will have its own biosafety regulations that should be adhered to, and anyone in any doubt about handling procedures should contact the local safety committee (and if there is not one, create it!). Rules and

regulations vary among institutions and countries, so it is difficult to generalize, but a good review can be obtained in Caputo [1996]. Genetic modification of organism can have unpredictable/ undesirable effects when such organisms are introduced into the ecosystem. The modification and use of such organism for public service has also resulted in problems with the granting of patents. Hence, the Indian Government has set up organizations which are authorized to make decisions regarding the validity of genetic modification and the safety of introducing genetically modified organisms fro public services. One such organization is the Genetic Engineering approval committee (GEAC).

Biosafety Levels

The regulations and recommendations for biosafety in the United States are contained in the document Biosafety in Microbiological and Biomedical Laboratories, prepared by the Centers for Disease Control (CDC) and the National Institutes of Health (NIH), and published by the U.S. Department of Health and Human Services. The document defines four ascending levels of containment, referred to as biosafety levels 1 through 4, and describes the microbiological practices, safety equipment, and facility safeguards for the corresponding level of risk associated with handling a particular agent.

Biosafety Level 1 (BSL-1)

BSL-1 is the basic level of protection common to most research and clinical laboratories, and is appropriate for agents that are not known to cause disease in normal, healthy humans.

Biosafety Level 2 (BSL-2)

BSL-2 is appropriate for moderate-risk agents known to cause human disease of varying severity by ingestion or through percutaneous or mucous membrane exposure. Most cell culture labs should be at least BSL-2, but the exact requirements depend upon the cell line used and

the type of work conducted.

Biosafety Level 3 (BSL-3)

BSL-3 is appropriate for indigenous or exotic agents with a known potential for aerosol transmission, and for agents that may cause serious and potentially lethal infections.

Biosafety Level 4 (BSL-4)

BSL-4 is appropriate for exotic agents that pose a high individual risk of life-threatening disease by infectious aerosols and for which no treatment is available. These agents are restricted to high containment laboratories. For more information about the biosafety level guidelines, refer to Biosafety in Microbiological and Biomedical Laboratories, 5th Edition, which is available for downloading at www.cdc.gov/od/ohs/biosfty/bmbl5/bmbl5toc.htm.

Biopiracy:

The industrialized/ developed nations are rich financially, bur poor in biodiversity and traditional knowledge, while the developing and underdeveloped countries are rich in bioresources and traditional knowledge. Some such developed countries use the bio resources and traditional knowledge of other countries without proper authorization and/ or compensation to the countries concerned (Biopiracy). Eg: Basmati rice grown in India is distinct for its unique flavor and aroma, but an American company got patent rights on Basmati through the US patent and trademark office; the new variety of Basmati has been developed by this company by crossing an Indian variety with semi-dwarf varieties. Now some nations are developing laws to prevent such unauthorized exploitation of their bioresources and traditional knowledge

Ethical issuesassociated with transgenics and cloning.

Technology isn't perfected yet with very low success rate, the animal developed has high mortality rates.

- **Safety/risk of consumption**
  According to the U.S. Food and Drug Administration cloned animals probably safe to raise and eat. Transgenic ones may not be safe to consume.
- **Animal welfare**
  Birth weights, longer gestation, difficult births in clones. Poor survival rate of fetuses using some techniques. Anatomical, physiological, behavioral abnormalities
- **Suffering of transgenic animals**
  Case of Beltsville pigs (human GH introduced): High mortality, arthritis, gastric ulcers, degenerative joint disease, infection, lethargy. Cloned animals found to have shortened life spans with health problems.

Implications for application of technologies to humans

Moral concerns: "are we playing God?" Impact on ecosystems and genetic diversity –What if GE organisms escape reproduce? What might be the impact of limited gene pools on livestock faced with new (deadly) pathogens? Potential for GE animals to move into areas previously unused for agriculture disrupt fragile ecosystems habitat preservation issues for wild animals.Lack of controls to prevent GE animals from entering the food chain (e.g., cows that produce drugs in their milk)

Animal biotechnology and law

"Any food system practice that does not allow individuals who do not want to consume meat or milk from clones to act upon their

values at a reasonable cost is ethically unacceptable and ought to be illegal." (Thompson, 1997)

Conclusions

Biotechnology has given to humans several useful products by using microbes, plant, animals and their metabolic machinery. Animal/Plant tissues can be dissociated into their component cells, from which individual cell types can be purified and used for biochemical analysis or for the establishment of cell cultures. Many animal and plant cells survive and proliferate in a culture dish if they are provided with a suitable medium containing nutrients and specific protein growth factors. Although many animal cells stop dividing after a finite number of cell divisions, cells that have been immortalized through spontaneous mutations or genetic manipulation can be maintained indefinitely in cell lines. Clones can be derived from a single ancestor cell, or by fusing two cells to produce heterocaryons with two nuclei, enabling interactions between the components of the original two cells to be examined. Heterocaryons eventually form hybrid cells with a single fused nucleus. One type of hybrid cell, called a hybridoma, is widely employed to produce unlimited quantities of uniform monoclonal antibodies, which are widely used to detect and purify cellular proteins. DNA technology has made it possible to engineer microbes, plants and animals such that they have novel capabilities. Genetically Modified Organisms have been created by using methods other than natural methods to transfer one or more genes from one organism to another, generally using techniques such as recombinant DNA technology. Since the recombinant therapeutics are identical to human proteins, they do not induce unwanted immunological responses and are free from risk of infection as was observed in case of similar products isolated from non-human sources. Transgenic animals are also used to understand how genes

contribute to the development of a disease by serving as models for human diseases, such as cancer, cystic fibrosis, rheumatoid arthritis and Alzheimer's. Gene therapy is the insertion of genes into an individual's cells and tissues to treat diseases especially hereditary diseases. Viruses that attack their hosts and introduce their genetic material into the host cell as part of their replication cycle are used as vectors to transfer healthy genes or more recently portions of genes.

Although animal production is being changed significantly by advances made in thousands of biotechnology laboratories around the world, benefits are reaching the developing world in only a few areas of conservation, animal improvement, healthcare and the augmentation of feed resources. Adopting biotechnology has resulted in distinct benefits in terms of animal improvement and economic returns.

References:

1. Bradley, A., Evans, M., Kaufman, M. H. and Robertson, E. (1984).Formation of germ-line chimeras from embryo-derived teratocarcinoma cell lines. Nature 309, 255-256.
2. Chambard, J.C., Franchi, A., Le Cam, A., and Pouyssegur, J, (1983). Growth factor-stimulated protein phosphorylation in G(/G i -arrested fibroblasts. J. Biol. Chem. 258, 1706-1713.
3. Chambard, J.C., Paris, S., l'Allemain, G., and Pouyssegur,J, (1987). Two growth factor signaling pathways in fibroblasts distinguished by pertussis toxin. Nature (London) 326, 800-803.
4. Freshney, R.I. (1993) Culture of Animal Cells, A Manual of Basic Technique, 3rd ed., New York: Wiley-Liss.
5. Freshney, R.I. (2002) Cell line provenance. Cytotechnology 39: 3–15.
6. Freshney, R.I. (2005) Culture of Animal Cells, a Manual of Basic Technique , 5th Ed. Hoboken
7. Hay, C.R., Baglin, T.P., Collins, P.W., Hill, F.G. & Keeling, D.M. (2000) The diagnosis and management of factor VIII and IX inhibitors: a guideline from the UK Haemophilia Centre Doctors' Organization (UKHCDO). British Journal of Haematology,111, 78–90
8. López-Casillas, F., Wrana, J. L. &Massagué, Betaglycan presents ligand to the TGFβ signaling receptor J. Cell73, 1435−1444 (1993)
9. Maas-Szabowski, N., Stark, H.-J.andFusenig, N. E. (2002). Cell interaction and epithelial differentiation.In Culture of Epithelial Cells (ed. R. I. Freshney and M. G. Freshney), pp. 31-63. New York: Wiley.
10. Nieman,D.C. (1994). Exercise, upper respiratory tract infection, and the immune system.Med. Sci. Sports Exerc. 26:128-139
11. Petra, M., de Jong, P.M., van Sterkenburg, A.J.A., Kempenaar, J.A., Dijkman, J.H., Ponec, M. (1993) Serial culturing of human bronchial epithelial cells derived from biopsics. In Vitro Cell Dev.Biol.29A:379-387.

12. Rexroad, C. E., Jr., Hammer, R.E., Behringer, R.R., Palmiter, R.D., and Brinster, R.L. (1990) Insertion, expression and physiology of growth-regulating genes in ruminants. J. Reprod. Fertile. 41, 119-124.
13. Schuman R.; Shoffner R. N. 1982.Potential genetic modifications in the chicken, Gallus domesticus. Proceedings of the 2nd world congress on genetics applied to livestock production 6: 157-163.
14. Stice, E., Schupak-Neuberg, E., Shaw, H.E., & Stein, R.I. (1994). Relation of media exposure to eating disorder symptomatology: An examination of mediating mechanisms. Journal of Abnormal Psychology, 103, 836–840.
15. Ward, S., and Klass, M. (1982).The location of the major protein in *Caenorhabditiselegans* sperm and spermatocytes. Dev. Biol. *92*, 203–208.
16. Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. Nature 171, 737–738.

Chapter 7
Medical Biotechnology

Harendra Modak, Madhuri Biradar, Nalina M
and Chandrashekara K.N.

Living cells and cell materials are used in the field of Medical biotechnology for research purposes to produce pharmaceutical and diagnostic products resulting in the treatment and prevention of human diseases. Insulin as well as growth hormones is the exemplary discoveries in the field of medical biotechnology. Tissue culture techniques are being used in detection of in born defects of metabolism and haemoglobinopathies; congenital abnormalities and understanding of clinical aspects of autosomal and sex chromosomal disorders. Medical biotechnology has advanced to develop organ culture, produce artificial blood and to perform transgenesis. The organ culture is performed to precisely model functions of an organ in numerous states and conditions by the usage of the existent *in vitro* organ itself. The substitutes for blood are under progress for transfusion in lieu of donor blood during emergencies or prolonged surgeries. Clinical trials are being conducted for the first generation of blood substitutes. Novel testing and screening measures have led to the donor blood supply progressively safe. Through transgenesis, an exogenous gene – called a transgene – is introduced into a living organism which leads to the exhibition of new properties by the organism and transference of those properties to its offspring. Various techniques viz. ballistic DNA injection, plasmid vectors, pronuclear injection, viral vectors and liposomes facilitates transgenesis.

The development in the biochemical diagnostic and screening tests are useful for a range of inherited metabolic disorders. Disorders like congenital disorders of glycosylation, lysosomal storage diseases, mucolipidoses and several inborn errors of metabolism are diagnosed by analyte and enzyme testing. Cancer cytogenetics with molecular aspects has undoubtedly made gigantic advances with the introduction and application of new methodologies. The introduction of FISH and other interrelated techniques viz. spectral karyotyping and comparative genome hybridization have furnished outcomes that could have been unknown otherwise. Whereas, Somatic cell genetics proved to be helpful in development of hybridoma technology followed by gene mapping and resultantly availability of polyclonal and monoclonal antibodies. Immunogenetics has provided a detailed knowledge about Immune deficiencies and disorders, blood groups and transplantation antigens, HLA and disease association followed by immunological diseases comprising AIDS, antigen processing and MHC. Dynamic mutations, caused by the expansion of unstable polymorphic DNA repeat sequences, can give rise to recessive, dominant or X-linked disorders, depending upon the position of the repeat sequence with respect to the genes that have received impact by the expansion. The characteristic feature of these mutations is the existing instability, which is a role of the copy number of repeats and can happen in either mitosis or meiosis. There is an association between repeat copy number and age-at-onset and/or severity of symptoms of the disease for several resultant disorders. Resultantly, a vast arena of research is now engrossed on recognizing the pathogenic ways from the mutation to the disease symptoms in the expectation of finding resources of delaying onset, slowing advancement or even preventing symptoms of the disease. The rising list of neuromuscular and neurodegenerative diseases caused by SBMA, HD, DRPLA, SCAs, OPMD, HDL-2, FRDA, FRAXA etc. Chromosome

walking as well as chromosomal localization of genes is used to detect a particular disease gene. The walking starts at the closest gene known as a marker gene and each successive gene in the sequence is verified continually for if any overlap restrictions and mapped for their accurate area in the sequence. When the gene is cloned, its function can be fully recognized. During the course of this process, assessments are done to fully categorize the properties of each successive clone, to map their positions for future use. G-banded chromosomal preparations are used for uncovering of autosomes of autosomal/sex chromosomal disorders (deletion, translocation, Klinefelter syndrome, Down's syndrome, Turner's syndrome, etc.); PCR bases diagnosis for. fragile-X syndrome and SRY in sex chromosomal anomalies); FISH for detections of translocations, inversions (using appropriate probes) (e.g., chro 9-22 translocation; X-Y translocation); DNA sequencing of representative clones to detect mutation(s); PCR-SSCP to detect mutations (e.g., sickle cell anemia, thalassemia);Southern blot-based diagnosis for trinucleoide expansions in fragile-X syndrome, SCA, etc.

History

Biotechnology has been applied in the medical science for hundreds of years with mankind's revelation that diseases can be cured from living organisms by using their products. The earliest known use of antibiotics can be traced back to 2500 years ago, when the mouldy curds made from soybeans were used by ancient Chinese to fight infection.

Louis Pasteur is considered one of the pioneers in the improvement of modern antibiotics. In 1870s, he discovered that saprophytic *bacillus* can check the growth of anthrax bacteria (*Bacillus anthracis*). Alexander Fleming discovered penicillinin in 1928.

In 1973, the medical age of biotechnology was started by Herb Boyer and Stanley Cohen when they could develop a technique of

introducing DNA into an *E. coli* bacterium and created a transgenic bacterium. Later this recombinant DNA technology was used to successfully introduce the human insulin gene into *E. coli*. The genetically engineered *E. coli* was able to synthesize human insulin. Based upon Boyer and Cohen's recombinant DNA technique, Werner Aber, Daniel Nathans and Hamilton Smith discovered restriction endonuclease enzymes and received Nobel Prize for Medicine in 1978.

1. Introduction

Biotechnology plays an important role in the interpretation of the molecular causes of disease and in the improvement of novel diagnostic methods and improved targeted drugs. This technology allows new methods to treat diseases with new module of drugs targeting unknown causes. Individual patients are also getting attention for their differences. The diagnosis and treatment of diseases have become increasingly interlinked. Diseases being diagnosed on the basis of molecular information can bring the opportunity of effective treatment which would depend mainly on the availability of diagnostic techniques. Patients are gained by the progress in medical biotechnology with more precise, safer and more satisfactory treatment of their illnesses.

Medical biotechnology offers novel therapeutic possibilities and lead to fresh ways of fighting diseases such as rheumatic diseases, cancer and diabetes. Early and unambiguous diagnosis, along with the tests monitoring treatment and the stages of an illness, result in a successful treatment of patients.

2. Background information

Since medical biotechnology being a vast field, some of the related topics have been described in this chapter. This chapter offers background information and actions in the following areas:

2.1. Human Genome Project and its influence on medical biotechnology

2.2 Detecting Genetic Diseases
2.3 Biotech in the Hospital
2.3.1 Pharmacogenomics
2.3.2 Gene therapy
2.3.3 Assays

2.1 Human Genome Project and its influence on medical biotechnology

The human genome project was started in 1990 under the coordination of the U.S. Department of Energy and the National Institutes of Health. The main goals of human genome project were to:

- identify all the approximately 25,000-30,000 genes in human DNA,
- determine the sequences of the 3 billion chemical base pairs that make up human genome

2.1.1 Applications of genome research

Some recent and budding applications of genome research include:

- Molecular medicine
- DNA forensics (identification in crime spots)
- Energy sources and environmental applications
- Agriculture, livestock breeding, and bioprocessing
- Bioarchaeology, anthropology, evolution, and human migration

2.1.2 Influence on medical biotechnology

Some of the goals of molecular medicine include:

- Improved diagnosis of disease
- Earlier detection of genetic predispositions to disease
- Rational drug design
- Gene therapy and control systems for drugs

- Pharmacogenomic (custom drugs)

Genome maps are progressively expanding which aid researchers in seeking genes connected with numbers of genetic diseases, including neurofibromatosis types 1 and 2, myotonic dystrophy, inherited colon cancer, fragile X syndrome, familial breast cancer and Alzheimer's disease. Additionally, a large no. of people dies every year from undesirable reactions to drugs, and others face painful or dangerous side effects. Now, it is known that genes and DNA sequences influence drug responses; it is anticipated that the number of adverse responses would reduce and side effects can be obviated.

2.2 Detecting Genetic Diseases

Earlier, genetic testing was usually performed on foetuses for the purpose of identifying the sex of a child or for identifying a number of genetic diseases such as syndromes. Mostly, amniocentesis was conducted in such cases and then karyotyping was performed. Karyotyping is also used for adults for checking missing, duplicate or defective chromosomes in adults. With the advancement of research, more refined techniques are being invented; and being useful to detect individual diseased genes in children and adults. Correct diagnosis of a genetic disorder allows for more rapid and effective application of appropriate treatment.

Gene Tests based on DNA can detect inherited breast and ovarian cancer, Amyotrophic lateral sclerosis (ALS), Hereditary nonpolyposis, Alzheimer's disease, Ataxia telangiectasia, Hemophilia A and B, colon cancer, Tay-Sachs Disease, Cystic fibrosis, Myotonic dystrophy, Spinal muscular atrophy, Sickle cell disease etc.

2.3 Biotech in the Hospital

2.3.1 Pharmacogenomics

It covers the aspect of designing of the most operative drug therapy and treatment approach based on the detailed genetic profile of a patient. Same drug produces different responses in different individuals. Pharmacogenomics gives hope that extensive genetic studies will lead to the production of personalized drugs and increase their safety and efficacy.

2.3.2 Gene Therapy

Briefly, Gene therapy is useful for correcting defective (missing or damaged) genes. This can be done by:

- Insertion of a normal gene into a nonspecific position
- Swapping of abnormal gene for normal gene
- Repairing an abnormal gene
- Turning a gene on or off

Typically, a carrier molecule called a vector is used for inserting a normal gene into the genome. A virus whose disease-encoding genes have been substituted by therapeutic genes is considered a common vector.

2.3.3 Assays and genetic testing devices

Rapid diagnostics technology, Immunodiagnostics, swiftly and efficiently detects proteins, antibodies and infectious agents in different range of formats within very few minutes with 100%accuracy rates. For example, the best known example of medical device made possible by biotechnology is a pregnancy test kit. In this home pregnancy test kit, monoclonal antibody (MAb), a protein, binds to HCG and causes a colour change. HCG is found in a woman's urine only during her pregnancy stage.

Several other biotech devices are offered to identify infections, drug levels, hormone levels, and cancer cells. These devices are also proved to be an important tool against infectious diseases and bioterrorism.

3. Diagnostics and Therapeutics

Biotechnology proposes a major avenue to developed therapies and treatments. The development and usage of biotechnology techniques have enabled us to define the structure of DNA as well as the coded message hidden in the gene of DNA. Evidently, the sequence of genes stretched on an organism's DNA is accountable for the individual traits.

It has been discovered that faulty genes cause few diseases of the human body viz. Huntington's disease, a degenerative disorder of nerve cells. A single gene that produces a defective protein results in nerve cell death.

The known genes for causing specific diseases can be used as targets to develop designed treatments for those diseases. Since the genes causing the disease is known, the finding of the treatment becomes cheaper and easier. The conventional pharmaceuticals are not able to reach these gene targets.

Another major reason is the large biological molecules like synthetic insulin for which biotechnology is being used. It cannot be manufactured in a traditional chemistry laboratory. Synthetic insulin can only be synthesized by living cells using biotechnological techniques. There are other large molecules also, produced by the means of biotechnology. For example, blood clotting factors for haemophiliacs, human growth hormone, fertility drugs etc.

3.1 Role of biotechnology in diagnostics

Biotechnology has a vital role to play in present diagnostic sciences. The modern knowledge of DNA and its genes can examine the DNA of a person for following aspects:

- Prenatal diagnostic screening for diseases viz. Down syndrome
- Determination of the sex of an unborn baby
- Screening of newborns for HIV, phenylketonuria
- Screening of carriers: the unaffected individuals are identified for carrying one copy of a gene for a disease, which actually requires two copies of a gene to produce the symptom, for example haemophilia and Tay Sachs disease
- Estimation of the risk for development of adult-onset cancers by presymptomatic testing: for example bladder cancer and colon cancer). A mutation in the BRCA1 gene in a person has a high risk of developing breast cancer
- Prediction of adult onset disorders by presymptomatic testing: for familial high blood cholesterol
- Confirmation of the disease in symptomatic individuals: for example Huntington's disease and cystic fibrosis
- Forensic testing: analysis of blood, semen, hairs for prosecution of offenders, paternity testing.

For diagnostic DNA testing, two techniques are generally used. The first technique involves comparing the sequence of bases in the gene of the patient to that of a healthy individual and in the second technique short pieces of DNA i.e. probes are used that consist sequences complementary to suspected mutations. These probes search for the complementary sequences among the patient's DNA, then bind to it and flag or mark it.

3.2 Role of biotechnology in therapeutics

Biotechnology helps in the development of treatments to cure diseases in mainly two ways: gene therapy and pharmacogenomics.

3.2.1 Gene Therapy

Gene therapy can treat diseases by changing the genetic information of cells. Even though, gene therapy is new aspect, cases are there when patients have been cured by this technique.

Gene therapy mostly follows one of these approaches:

- the substitution of a mutated gene with a normal gene
- the knocking out the activation of the mutated gene
- Insertion of a completely new gene into the cells.

At present, somatic gene therapy is performed on body cells. In this therapy, changes in genetic information's are not passed on to the offspring's. By contrast, in Germline therapy, the reproductive cells (sperm and ova) are involved and any changes in genetic information's will be transferred on to the progenies. This type of gene therapy is still in its infancy and is not being experimented on patients.

In somatic gene therapy, the vector virus delivering the gene of interest is introduced into the patient intravenously. Consequently, new gene is delivered into the cell as the vector infects the target cells.

In 1990, the first case of gene therapy was performed by doctors at the National Institutes of Health (USA) for treating a toddler girl suffering from SCID (severe combined immune deficiency). SCID is caused due to mutation in ADA gene, which regulates the formation of an enzyme, adenosine deaminase. The medical practitioners took out bone marrow cells from the girl and treated it with a vector inserted with a normal ADA gene; and then re-introduced the treated bone marrow cells into the girl. The treatment made her immune system to start functioning normally.

Gene therapy is also being explored by scientists to treat cancer cells. Several possibilities like replacing altered or missing genes causing cancer or introduction of new genes into cancer cells are being experimented to make them more susceptible to treatment.

3.2.2 Pharmacogenomics

Pharmacogenomics would help the doctors to recommend the correct medication and dosage on the basis of patient's genetic profile, thereby minimizing the risk of undesirable reactions, over dosage and side effects. Identification of SNPs is the basis of pharmacogenomics. Earlier, the sequencing of a genome was an expensive and lengthy procedure, but with the improvement of the DNA microarray, the sequencing has become easier and quicker. SNPs are being used to map and recognize specific genes that lead to the development of diseases such as cancer, arthritis and diabetes. The proteins produced by these genes become targets for novel therapies.

Pharmacogenomics play a significant part in treatment of blood cholesterol levels, oncology, and treatment for patients with cardiovascular diseases, tailoring treatment for patients with psychiatric disorders.

Nevertheless, Pharmacogenomics is still in its infancy to individualise treatment procedures. Pharmacogenomic testing to decide a patient's potential reaction against a treatment is sophisticated and advanced in only few countries.

4. Role of biotechnology in developing vaccines

A vaccine is a safe biological preparation, administered to humans to make them immune against a particular disease. The immune system of the human body identifies the vaccine as a foreign particle and destroys it, but the memory of the foreign matter remains intact with the immune system. Later, when the person factually gets infected with

the virulent form of the disease causing organism, the immune system immediately recognises it and fight against the infection by releasing antibodies.

There are several possible ways to develop a vaccine and it depends upon the procedure of infection by disease-causing microbes, the response of the immune system, the target site of the vaccine and the different physical characteristics of the microbe. Monovalent vaccines can immunize against one disease and multivalent vaccines immunize against two or more strains of the same disease-causing microbe or more than one disease.

4.1 The various approaches are as follows:

4.1.1 Live, attenuated vaccines

These types of vaccines comprise attenuated form of a disease-causing microbe, so that it can no longer cause disease but only stimulate the immune system to memorize it. These vaccines produce a strong immune reaction and only one or two doses of the vaccines are enough to give immunity for whole life. It is predominantly used against viral diseases viz. chickenpox, mumps and measles. It should not be administered upon a person having a weak immune system such as HIV patients and patients undergoing chemotherapy. Some of the drawbacks are that they need to be refrigerated to remain effective which limits their utilization in some poor countries. Rarely, the weakened microbe can revert back to its virulent form.

4.1.2 Inactivated vaccines

In these types of vaccines, microbes are not weakened but inactivated by killing them using radiation, heat or chemicals. Therefore, the microbe is unable to mutate back to its potent form. Booster doses are required as this kind of vaccine does not generate a strong immune response. These can be used in developing countries since they are

freeze-dried, thus, are stored easily. For example, inactivated vaccines are manufactured against cholera, hepatitis A and bubonic plague.

4.1.3 Toxoid vaccines

To combat diseases like tetanus and diphtheria caused by toxin secreting bacteria, these vaccines are used. It is made by exposing the toxin to formalin and making the toxin harmless. The immune system releases antibodies against the toxin which bind and block the action of toxins.

4.1.4 Subunit vaccines

These vaccines contain only the antigens of disease-causing microbe and not the entire microbe. Evidently, antigens induce the immune system the most. Antigens are recognised and bound by the T-cells of the immune system. Using recombinant DNA technology antigens can be made from the microbes in the laboratory. Such vaccines are referred as combining subunit vaccine. The vaccine against the hepatitis B virus is one of the examples.

4.1.5 DNA vaccines

Although vaccines against influenza and herpes are currently being administered, these types of vaccines are still in the trial phase. In DNA vaccines, when the genes that code for the antigens of that microbe are introduced into the body cells they initiate the body cells to generate antigens. As a result, the antigens stimulate the immune response of the body. These vaccines are easily producible and storable.

4.1.6 Conjugate vaccines

In bacteria with polysaccharides outer coating cannot be recognised by the immature immune system of a baby as polysaccharides mask the antigens present on the surface of bacteria. These antigens from the microbe are linked to the polysaccharides, so that the immature immune system can recognise them. These are known

as conjugate vaccines. *Haemophilus influenzae* type B (*Hib*) is one of the examples for conjugate vaccine.

4.1.7 Recombinant vector vaccines

Functionally similar to DNA vaccines, recombinant vector vaccines use a different method to introduce itself. The vector, either an attenuated bacterium or virus, is used to carry the DNA of the microbe into the body of the patient and infects it followed by delivering the DNA to the body cells. Research is in progress to develop bacterium-based and viral-based recombinant vector vaccines against rabies, HIV and measles.

5. Treatments developed with the aid of biotechnology

5.1 Treatment of Cancers

Treatments using monoclonal antibodies are being used now-a-days against numerous forms of cancer. In this type of treatment, synthetic monoclonal antibody is attached with the cancer cell followed by killing the cancerous cells in different ways. The monoclonal antibody Cetuximab blocks the growth signals produced by the cancer cells resulting in the halting of the cells growth and thus colon cancer is treated. Gemtuzumab combined with strong chemotherapeuticals are administered into the cancer cell where they become active and minimize damage to adjacent normal tissues, help in treating acute myelogenous leukaemia (AML). Rituximab facilitates the cancer cell to be more visible to the immune system to get destroyed. This antibody is useful in treating non-Hodgkin's lymphoma. Herceptin is a monoclonal antibody to treat breast cancer cells in women expressing the protein HER2. Herceptin specially binds to those cancer cells and discontinue their proliferation. A radioactive particle is combined with Ibritumomab monoclonal antibody to deliver the radiation directly to the cancer cells

which does not harm the neighbouring normal tissues. Non-Hodgkin's lymphoma is treated in this way also other than Rituximab antibody.

5.2 Blood clotting factors for haemophiliacs

Generally, haemophilic patients must go through regular infusions of the missing clotting factor VIII (most common form) to treat the disorder. In some cases; patients develop antibodies against the infused clotting factor, making the replacement unsuccessful. VIIa, a recombinant human factor, has been synthesized to effectively treat intense bleeding in patients by circulating inhibitors.

5.3 Bone marrow transplantation

Since 1950s, bone marrow transplants are being performed to treat the patients suffering from different disorders of the blood viz. leukaemia. The patient undergoes chemotherapy to kill cancerous cells before a transplant. To get an exact match, bone marrow is transplanted either from the patient's own body or from a person having matching genetic make-up. To check the genetic make-up similarity, blood samples from donors are examined to know their HLAs (human leukocyte antigens).The donor's bone marrow cells may attack the patient's tissue if the donor and patient do not have matching HLAs.

5.4 Xenotransplantation

Due to the chronic shortage of donors for various organs or tissues, biotechnology is trying to solve this problem by xenotransplantation. Here the donor organs are procured from different species like monkey and pigs. Although this technology is still being used on a small scale, nearly 60,000 transplants of heart valves, procuring heart valves from pigs, are accomplished in the USA every year.

5.5 Preventing risk of rejection after organ transplants

Immunosuppressant medication is used by doctors to inhibit donor organs being rejected by the patient's body. Albeit this prevents the organ from being rejected, it weakens the patients' immune system

and enhances the susceptibility to different infections. Biotechnologically derived, Cyclosporine, available naturally from a fungus that grows in soil, suppresses only a part of the immune system involving rejection and produces a less severe effect on the remaining part of the immune system.

6. Other Applications of Biotechnology in Medical field

6.1 Medical Applications of Molecular and Cellular Immobilization Techniques

Immobilization technique is applied in various ways in medicine. One of the applications is related with the regulation of equilibrium between coagulation and dissolution of coagulated blood i.e. fibrinolysis. Fibrinolytic therapy is used by the doctors for treating occlusions in some parts of the body where surgery is risky thus preventing deaths caused by thrombosis. Heparin and Plasmin are important enzymes having been immobilized for employ in such therapy.

6.2 Genetic Manipulation of Microorganisms

The performance of microorganisms in the medical, pharmaceutical and industrial application has been enhanced by genetic manipulation technique. Some generic techniques have been described briefly with their applications. These techniques are used to increase the performance of cells or microorganisms for particular applications.

6.2.1. Mutagenesis-based Techniques

Mutagenesis brings changes in a DNA sequence and alters either the structure of the gene products or the expression of genes. It is sub categorized into two types based upon the type genetic manipulation: classical mutagenesis and transposon-directed mutagenesis.

Chemical mutagens are used in classical mutagenesis to control required processes in the target microorganism. The mechanism of

action present in the targeted organism is traced through this classical approach.

In transposon-directed mutagenesis, the function of specific site of chromosomes is determined through genetic alteration of particular site on the target chromosome. This mutagenesis is more specific and used in cases where higher selectivity is required.

6.2.2. Gene Transfer

In electroporation, the induced transient pores in a bacterial cell membrane allow the introduction of DNA inside the cell under preset laboratory conditions. Such methods are used in various microorganisms and even with plasmids that are of major use in the pharmaceutical industry.

The transference of genes can also be performed by Shuttle vectors. These vectors are actually such DNA constructs which are able to replicate and deliver DNA to usually different types of bacteria (Blaschek, 1996). The mechanism of action of the microorganism is thoroughly studied to perform DNA transference, for example E. coli strains.

6.3 Microencapsulation Techniques and their Applications

Microencapsulation is used in the production of enzymes, for the transplant of organs and manufacturing capsules for the production of vaccines. The prospects of microencapsulation techniques can be used with other new biotechnology tools to expand the range of its applications.

Several classical techniques used for microencapsulation are spray drying, molecular inclusion and extrusion. Liposome microencapsulation is an advanced technique that can introduce substance of interest in an organism regardless of certain aspects like

electrical charge, the substance's solubility, molecular size or structure (Gregoriadis, 1984).

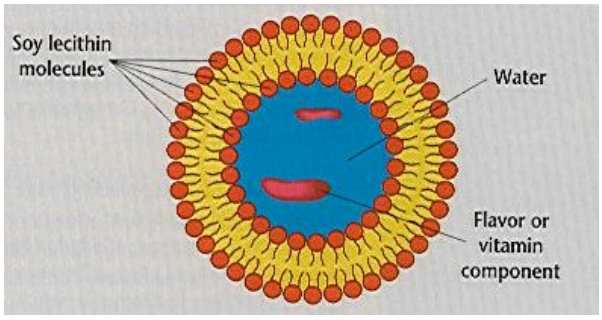


Figure 1: Simplified representation of the molecular organization of a liposome microcapsule in water.

6.3.1. Liposome Applications in Medical Biotechnology

Conventionally, liposomes (Fig.1) are used in the bioengineering field to generate certain proteins by genetic modification of cells. This has solved the difficulty of transferring high molecular weight molecules through cell membranes (Nicolau and Cudd, 1989).

Vaccine preparations based on liposomes have been lucratively tested in immunization of animals and such experiments are presently in the clinical testing phase. The advantages and disadvantages of liposomes as drug carriers in a human body system depend basically on the immuno-compatibility, interactions of liposomes with the cells (Lasic, 1993), or their capacity to avoid detection by the human immune system.

6.4 Lectin Applications to Cancer Detection and Treatment

Lectins are bioactive proteins and glycoproteins that agglutinate erythrocytes of some or all blood groups *in vitro* (Sharon, 1998). They are

found in most organisms, including bacteria, viruses, invertebrates, vertebrates and plants.

Plants have several phytochemicals that can change biochemical pathways associated with cancers. One of such phytochemicals is lectins which are being deeply examined for its role in cancer chemoprevention (Abdullaev and Gonzalez de Mejia, 1997). Lectins bind to cancer cell membranes or their receptors causing cytotoxicity, tumor growth inhibition and apoptosis. They also affect the immune system by activating certain protein kinases or by changing the production of various interleukins (Gonzalez de Mejia and Prisecaru, 2003).

Lectins are also reported to have inhibitory effect on the growth of tumors. Their potential for clinical applications has been investigated only in recent years. Lectins are applicable in the diagnostics as well as therapeutics of cancers. Thus, lectins are considered to be versatile biomarkers and have been utilized in histochemical, biochemical and functional studies of cancer cell characterization (Munoz *et al*, 2001).

6.5 Applications of Molecular Modelling

A group of techniques together make Molecular modellingand use computer-generated images of different chemical structures to understand physicochemical properties of molecules, to know the relative positioning of atoms present in the molecule and to offer clues about their potential roles i.e. structure-functionrelationships in the organism. Therefore, when the structure for one member of a protein family is understood, molecular modelling calculations or homology technique can help identify the structure for other members of the same protein family as proteins of same family share similar sequences and the same basic structure.

Molecular modelling is useful in medical fields, for example, in the development of new drugs on a nanoscale. The molecular modelling

of Epigallocatechin Gallate (EGCG) and the HIV cells was undertaken by Shearer (2003). His report has inspired scientists in Japan who discovered the potential of green tea as an anti-HIV drug. *Epigallocatechin Gallate* (EGCG) is a chemical compound, present abundantly in the green tea, is found to impede the HIV virus from binding to CD4 molecules and human T-cells so EGCG is being tried as an anti-HIV drug.

6.6 Designer Babies

6.6.1 Introduction

The phrase designer babies generally indicate the genetic interventions into the pre-implantation embryos in the effort to manipulate the traits the desired children will have. The term "Designer babies" is a journalistic term used by commentators, not by scientists, to explain different reproductive technologies. Progress in three different fields has made designer babies possible:

6.6.1.1 Advanced Reproductive Technologies

The examples of such technologies are:

a. *In vitro* fertilization
b. Frozen embryos
c. Direct injection of a sperm cell into an egg
d. Egg and sperm donations
e. Pregnancies by older women
f. Surrogate motherhood
g. Cell and Chromosome Manipulation
h. Genetics and Genomics

6.6.2 Arguments for Designer babies:

- It helps to prevent certain genetic diseases in children to protect them from suffering debilitating diseases and deformities, thus minimizing the financial and emotional strain on the parents

- By screening embryos, naturally conceived but defected embryos are rejected from the womb

6.6.3 Arguments against Designer babies

- In the course of making designer babies, embryos are terminated which does not fit with moral and religious criteria.
- Bias against people already born with physical or mental disabilities would raise and they might be perceived as genetically inferior
- There is always the threatening shade of eugenics. Designing of babies would give us venues to present our genetic preferences which might be dangerous in future.
- The tampering of human genetic structure could damage the gene pool leading to undesired and unpredictable consequences.

7. Miscellaneous: New Biotech Breakthroughs about to change Medicine

The recent advances in the world of medicine have blurred the differences between biology and technology. Some of these examples are:

1) **Decay-Fighting Microbes:** Since, bacteria living on teeth transform sugar into lactic acid, which erodes the enamel and lead to tooth decay, a new bacterial strain, called SMaRT, has been bioengineered by Florida-based Company ONI BioPharma. The antibiotic released by SMaRT kills the natural decay-causing strain. One swabbing of SMaRT onto teeth would keep them healthy for a whole life.
2) **Artificial Lymph Nodes:** Doctors could fill the nodes with artificial versions of lymph node cells explicitly tuned to treat certain conditions, such as HIV or cancer.
3) **Asthma Sensor:** Asthma accounts for a quarter of all emergency room visits in the U.S., but a sensor developed at the University of

Pittsburgh may finally cause that number to plummet. A handheld sensor is developed which is a polymer-coated carbon nanotube and 100,000 times thinner than a human hair. It analyzes breath for small amounts of nitric oxide as this gas is produced in prior to asthma attacks.

4) **Cancer Spit Test:** Researchers at the University of California, Los Angeles detect oral cancer from a single drop of saliva. Proteins associated with cancer cells respond with dyes on the sensor and emit fluorescent light, detectable by microscope.
5) **Prosthetic Feedback:** The prosthetic limbs are difficult to monitor because they can't sense the position of limbs. Stanford University graduate student Karlin Bark is developing a device. It stretches the skin of amputee near the prosthesis in such a way that it provides feedback about the limb's position and movement.
6) **Speech Restorer:** Illinois-based Ambient Corporation has developed a device called phonetic speech engine for patients who have lost the ability to talk. Here, patients imagine slowly sounding out words; then this quarter-size device, located in a neck brace, transmits those impulses wirelessly to a computer or mobile phone and produces speech.
7) **Muscle Stimulator:** In the time it takes for broken bones to heal, nearby muscles often atrophy from lack of use. Israeli company StimuHeal has developed MyoSpare, a battery-operated device. It uses electrical stimulators to solve the problem of atrophy of muscles nearby broken bones. It facilitates the exercise of muscles and maintains them strong during recovery.
8) **Nerve Regenerator:** A nanogel developed at Northwestern University eliminates that impediment of growth of nerve fibres along the injured spinal cords.

9) **Rocket-Powered Arm:** Vanderbilt University scientist Michael Goldfarb developed a power source: rocket propellant in prosthetic arm which can lift 20 pounds, three to four times more than current prosthetics. This power source is pencil-size version of the mono-propellant rocket-motor system used to move the space shuttle in orbit. H_2O_2 powers the prosthetic arm for 18 hours to perform normal activity. This has the potential to replace bulky battery packs required to add strength to prosthetic limbs.

References:

1. Blaschek HP. 1996. Recent Develpoments in the Genetic Manipulation of Microorganims for biotechnology applications. In: Baianu IC, Pessen H and Kumosinski TF, eds. *Physical Chemistry of Food Processes*. Vol 2. New York: Van Nostrand Reinhold. p. 459-474.
2. Gregoriadis G, ed.1984. *Liposome Technology.* New York: CRC Press.
3. Nicolau C. and Cudd A. 1989. Liposomes as carriers of DNA. *Crit. Rev. Therap*. *Drug Carrier Systems.* 6: 239–271.
4. Lasic DD. 1993. Liposomes: *From Physics to Applications*. Amsterdam: Elsevier.
5. Sharon N. 1998.Glycoproteins now and then: a personal account. *Acta Anat* (Basel). 161(1-4):7-17.
6. Abdullaev FI, de Mejía EG. 1997. Antitumor effect of plant lectins. *Natural Toxins*. 5:157-163.
7. Gonzalez de Mejia E, and Prisecaru VI, Lectins as Bioactive Plant Proteins: New Frontiers in Cancer Treatment. 2005. *Crit. Rev. Food Sci. & Nutr*. 45: 425-445.
8. Munoz R, Arias Y, Ferreras JM, Jimenez P, Rojo MA, and Girbes T. 2001. Sensitivity of cancer cell lines to the novel non-toxic type 2 ribosome-inactivating protein nigrin b. *CancerLett.* 167(2): 163-169.
9. Nance CL, Shearer WT.2003. Is green tea good for HIV-1 infection?. *J Allergy Clin Immunol*.112:851–853.

Chapter 8
Tools and Techniques in Biotechnology

Harendra Modak, Chidanand Rabinal, Madhuri Biradar,
Nalina M and Chandrashekara K.N.

The fundamental unit of cell, genetic material, decides the type of cell and its behavior. The study with reference to nucleic acid, expression pattern of genes, proteins and other final products, formation of metabolites and their types gives a reflection about the cell and its function. Advancement in biotechnological methods has made these studies approachable and applicable. Since, biotechnology is the application of biological organisms, system or processes to manufacturing and service industries related to animals, plants, microbes and environment, here we are exploring the diverse tools and techniques involved in genomics, transcriptomic and proteomics. The genomic approaches facilitate to study the genetic material, locate the position of genes, chromosome paintings and hybridization techniques. The expression pattern of these genes studied by RNA isolation, qRT PCR, differential display techniques and recent methods like SAGE, MPSS and RNA sequencing provide in depth study of gene expression. Additionally, protein studies like SDS-PAGE, 2D gel electrophoresis are described in this chapter.

BASIC TOOLS IN THE BIOTECHNOLOGY

Biotechnology is an interdisciplinary science that borrows scientific instruments commonly used in chemistry, biochemistry, genetics, and physics laboratories. Very few instruments are specifically

designed for biotechnology. Those that are unique to biotechnology were developed for the specific needs of particular research studies. A trip to a biotechnology laboratory would seem very much like a visit to any other science laboratory. This is also true for large facilities that produce biotechnology products. The machinery is used in many other industries. However, biotechnology instruments are focused on analyzing, manipulating, or manufacturing the chemicals that make up organisms. The major chemicals of interest in biotechnology are biological molecules called nucleic acids and proteins. Each instrument mentioned in this chapter can be found in most biotechnology industrial settings. Research laboratories are usually limited to particular equipment for research being performed.

Most of the tools of biotechnology are used to identify and isolate many of the biological molecules making up an organism. The identification of biological molecules is called characterization. Characterization tells researchers the specific chemical makeup of a molecule. General chemical characterization techniques help scientists in identifying molecules as one of four major biological molecule categories: carbohydrates, lipids, proteins, or nucleic acids. Resolution is a term used to describe the degree of detail used to characterize molecules. For example, high-resolution characterization provides information about the specific identity of a particular type of biological molecule. Many of the tools described in the following section tell researchers whether a particular protein or sequence of nucleic acids is present in a sample. Isolation is a method of separating a particular molecule from a mixture. Researchers interested in working with a pure sample of a molecule must isolate and collect it from a mixture. Many of the tools that identify molecules also isolate that molecule from the mixture, saving the researcher time and effort.

The first biotechnology tools date back to fermentation jars used to make alcoholic beverages used by ancient people almost 7,000 years ago. Special ceramic pots designed to enhance fermentation were discovered in archeological sites throughout Asia, the Middle East, and South America. Almost 3,000 years ago the Chinese were using devices for culturing and extracting antibiotic chemicals from moldy soybean curd. A boom in scientific instruments started in Europe after the 1600s with the advent of the microscope and new apparatus for conducting chemical reactions. The harnessing of electricity to operate machines refined the instruments used in older biotechnology applications. In addition, electricity permitted scientists to develop the great variety of analytic instruments used every day in biotechnology. By the late 1800s many of the instruments such as centrifuges and incubators seen in modern biotechnology laboratories were being developed. Improvements in electrical circuitry, motors, and robotics further refined the types of instruments used in biotechnology. Instruments were becoming more accurate and simpler to use. The advent of computers fuelled tremendous improvements in biotechnology instruments. Almost all of the instruments used in biotechnology today have a built-in computer or are linked to computers that integrate the instrument with other tools of biotechnology. Computers also make it possible to replace chart paper and older ways of collecting and recording data. This data can now be imported into other instruments or into software that carries out various types of analyses and statistical calculations. The computer can also place the data into an electronic notebook that could be e-mailed to other scientists.

Advances in miniaturization and the creation of lightweight materials for constructing instruments are providing new directions in biotechnology instrument design. Instruments that at one time took up all of the space on a laboratory table can now fit into an area of the size

of a small toaster. Portable instruments are making it possible for scientists to share and transport expensive and specialized instruments. This is particularly important in bioprocessing operations in which it is favourable to carry out instrumentation procedures at difficult locations of a facility. Miniaturization is leading to the development of microscopic instruments that can be placed into cell cultures of whole organisms for continuous monitoring. New methods of wireless communication are enhancing the ability of the instruments to transfer data. Scientists now have access to instruments that use devices similar to cell phones that can control instruments and transmit data to various computers.

We will explain how to use, calibrate and troubleshoot many pieces of equipments used in biotechnology labs. Some of them are described here based on their function:

1. **Measurement of Volume**
   - **A) Erlenmeyer flasks** – Erlenmeyer flasks are used primarily to prepare solutions prior to an accurate volume adjustment. Although there are volumetric markings on these flasks, they are not calibrated and should not be relied upon for exact volume measurements.
   - B) **Beakers** – Beakers are also used for preparing solutions, especially when a pH adjustment requires access to the solution by a pH probe. The volumetric markings on beakers are also not reliable.
   - C) **Graduated cylinders** – Graduated cylinders are calibrated with sufficient accuracy for most volume measurements when preparing solutions. For example, the calibration of most 100 mL graduated cylinder can be relied upon to accurately measure to within +/- 0.6 mL.
   - D) **Volumetric flasks** – Volumetric flasks are used to measure a specific volume with the highest degree of accuracy, and are

used to make standard solutions for analytical assays. For example, the calibration of a 100 mL volumetric flask can have an accuracy of +/- 0.1 mL.

E) **Pipets** – Pipets are glass or plastic devices that are routinely used to measure and transfer liquids by drawing the liquid into the tube with a bulb or mechanical pump.

a. Pasteur pipets are small glass tubes used with a bulb to transfer volumes as small as a single drop and as large as a few mL. They are not graduated and are not used to measure volumes.
b. Beral pipets (transfer pipets) are plastic pipettes with a bulb at one end used for transfer of liquids. Sometimes they have calibration marks, which have a low level of accuracy. They are often disposable, sterile and individually wrapped.
c. Serological, or "blowout," pipets are graduated glass tubes used to measure anywhere from 0.1 to 50 mL. When the liquid has drained from this pipette, the final drop in the tip is transferred with a puff of air.
d. Mohr, or "to deliver," pipets are similar to blowout pipettes, but do not require a puff of air to accurately deliver the desired volume. They can be identified by the label — "TD" on the top.
e. Volumetric pipets are not graduated, but are carefully calibrated to deliver a single, highly accurate volume, and are used for the transfer of exact volumes.
f. Automatic micropipettes are mechanical pumps calibrated to deliver highly accurate volumes generally less than 1.0 mL, and as little as 0.1 microliters. They are often adjustable for measuring different volumes and they always use dispensable plastic tips to actually transfer the liquids.

g. Multichannel micropipettes can deliver the same volume from as many as 12 tips simultaneously. All automatic micropipettes need regular maintenance, calibration, and validation.

2. Measurement of Weight

Instruments for weighing materials are called balances, and most laboratories have more than one type of balance, depending on the amount of material being measured and the degree of accuracy required.

a. **Mechanical balances –** Mechanical balances weigh an object on a pan hanging from a beam that has a counterbalanced weight. We do not use mechanical balances in our lab.

b. **Electronic balances –** Electronic balances have replaced most mechanical balances due to their greater accuracy and ease of operation. They are easier to use because they usually have a digital readout, and weighing dishes can be tared to read zero mass before using. Most balances used for preparation of solutions have a sensitivity of +/- 0.01 g, but electronic analytical balances can be sensitive to +/- 0.1 mg or less. Electronic balances require routine maintenance and recalibration.

3. Measurement of pH

Most solutions prepared in the biological laboratory must have a carefully controlled pH. Buffers are prepared by adjustment to a specific pH with strong acid and base solutions, using a meter to monitor the pH. A pH meter is a volt meter that measures the electrical potential between two electrodes. One electrode is in contact with your solution, and the other is in contact with a reference solution. Usually both of these electrodes are combined in a single pH probe that you place in your solution. These meters can read to the nearest 0.1 pH unit, but require frequent calibration with reference buffers of known pH.

4. Measurement of light

Solutions are often analyzed in the biotechnology lab by measuring how the solutes interact with light. A spectrophotometer measures the amount of light that is absorbed by a solution at a specific wavelength or over a range of wavelengths. If you know a wavelength at which a specific substance absorbs light, you can calculate the amount of that substance in a solution from the measured absorbance of that solution at that wavelength.

A. **VIS spectrophotometer** – A visible (VIS) spectrophotometer measures absorbance of light in the visible region of the spectrum (wavelength of about 400-700 nm). A small vessel called a cuvette, which is generally plastic or glass and which usually has an internal diameter of 1.0 cm, is filled with the solution and placed in the spectrophotometer for measurements.

B. **UV/VIS spectrophotometer** – An ultraviolet/visible (UV/VIS) spectrophotometer can also measure absorbance of light in the ultraviolet region of the spectrum (about 100-400nm). These spectrophotemeters require a halogen light bulb that emits ultraviolet light and require special cuvettes that don‘t absorb UV light.

C. **Scanning spectrophotometer** – A scanning spectrophotometer can measure the absorbance of a solution over a range of wavelengths, creating an absorbance spectrum that can be used to identify substances in a solution.

D. **NanoDrop spectrophotometer** – A NanoDrop spectrophotometer is a brand of scanning UV/VIS spectrophotometer that allows the user to measure the absorbance of a very small sample of liquid (1-2 uL). This instrument makes it easy to quickly evaluate the quality and quantity of nucleic acids or proteins in a small sample prep.

E. **Microplate reader** – A microplate reader is a spectrophotometer that can measure the absorbance in the individual wells of a plate. Usually the plates are 96 wells, but other formats are available, such as 48 wells and 384 wells. This allows the user to prepare and read many small samples at once, saving time and money. Microplate readers may also be capable of reading fluorescence and chemiluminescence, which are two types of light emission that are frequently used in biological research.

5. Solution Preparation

Solution preparation involves mixing liquids and dissolving solids in liquids. There are many specialized devices in addition to balances, volume measuring devices, and pH meters involved in these processes.

A. **Magnetic stirrers** – Magnetic stirrers come in the form of a box with a magnet inside attached to a motor that spins the magnet. When a vessel containing a magnetic stir bar is on top of the magnetic stirrer, the stir bar spins and stirs the contents of the vessel.

B. **Vortex mixer** – A vortex mixer rotates the bottom of a tube rapidly; setting up a vortex in the liquid that rapidly mixes the contents.

6. Microbiological techniques

Specialized equipment is required to isolate, transfer, and grow up cultures of microbes and tissues in the laboratory.

A. **Autoclaves** – Autoclaves are machines that achieve a high internal temperature and pressure and are used to sterilize solutions and glassware. The kitchen pressure cooker achieves the same results and can be used instead of an autoclave.

B. **Cell Culture Hood** – A biological safety or cell culture hood filters small particles out of the air in order to avoid contamination of cultures or sterile media. The filters are similar to those used to decontaminate air for operating rooms in hospitals or clean rooms used in the semiconductor industry.

C. **Fermentors** – Fermentors are used to grow up a large quantity of cells with automatically controlled pH and levels of oxygen and other nutrients.

D. Since most cells are generally too small to be seen with the naked eye, microscopes are used to magnify their images. Light or Bright field microscopes and inverted microscopes are the most common types found in biotechnology laboratories.

7. Preparation of biological samples for analysis

There are many pieces of equipment that are used to prepare biological samples for analysis.

A. **Sorvall-type centrifuge** – A Sorvall-type centrifuge, or preparative centrifuge, has a balanced rotor that holds vessels and spins them at high speed, up to 20,000 rpm. This will cause most insoluble particles such as cells and many subcellular components to rapidly form a pellet at the bottom of the vessel. Rotors are available that hold vessels as small as a few milliliters to as large as a liter. These centrifuges are often refrigerated so that heat-sensitive compounds are not damaged during centrifugation. We do not have one of these in our lab.

B. **Tabletop centrifuge** – A tabletop, or clinical, centrifuge is generally not refrigerated and spins at a much slower speed than a preparative centrifuge. Rotors for clinical centrifuges generally hold tubes with a capacity of 15 mL or less.

C. **Microcentrifuge** – A microcentrifuge holds microcentrifuge tubes that can hold about 1.5 mL of liquid. These microcentrifuges can also spin at high speeds and are sometimes refrigerated.

D. **Sonicator** – A sonicator emits ultrasonic waves that can be used to disrupt cells, allowing their contents to be released into the surrounding buffer in ―grind and find‖ strategies.

8. Separation of macromolecules

Since there are thousands of different macromolecules in each cell, purification of a specific one from all the others requires powerful separation techniques, such as chromatography and electrophoresis. Both of these approaches take advantage of physical and chemical properties that differ between the individual macromolecules.

A. **Gel electrophoresis** – In gel electrophoresis, the macromolecules are placed in a solid matrix, called a gel, which is under a liquid buffer. An electric field is applied to this system, and since biological macromolecules carry ionic charges, they will be attracted towards one pole of the electric field and repelled by the opposite. Thus, macromolecules characteristically migrate in either direction in the field. The migration speed is determined by the charge-to-mass ratio of the macromolecule.
   - In a flat gel, also called a horizontal or submarine gel, electrophoresis system, an agarose gel lies horizontally below the electrophoresis buffer. This technique is mainly used to separate large nucleic acids (DNA and RNA).
   - A vertical electrophoresis system holds a polyacrylamide gel in the vertical position, and is mainly used to separate proteins or small-sized nucleic acids.

B. **Chromatography** – Chromatography is a family of methods used to separate macromolecules through their relative affinity to a stationary phase (generally, solid chromatography beads) and a mobile phase (generally, an aqueous buffer). The chromatography beads are loaded into a tube, called a chromatography column, and buffer is dripped, or pumped, through the column to carry the macromolecules along. The macromolecules separate on their affinity for the mobile front. Some chromatography beads separate by charge (ion exchange chromatography), by hydrophobicity

(hydrophobic interaction chromatography), or by a specific property of that protein (affinity chromatography). Macromolecules can also be separated by size otherwise known as size exclusion or gel filtration chromatography.

- To overcome this limitation, high performance (or high pressure) chromatography (HPLC) uses high-pressure pumps and metal-jacketed columns to operate at high pressures and speed up the process.
- A fraction collector collects the released mobile phase (eluent) of a chromatography column. It automatically measures a programmed volume (sometimes by the number of drops of liquid) into a line of test tubes or microcentrifuge tubes.

9. Manipulation of Nucleic Acids

To analyze DNA some of the specialized pieces of equipment are:

A. **Thermal cycler** – A thermal cycler is a machine that is used for amplification of a specific section of DNA by PCR (polymerase chain reaction). The machine cycles through several temperatures, which allows an enzyme called DNA polymerase to use chemicals in solution to build DNA molecules identical to a template provided.

B. **Electroporator** – An electroporator is used to discharge a high-voltage, high-amperage pulse of electricity of very short duration through a cuvette containing suspended cells to disrupt their plasma membranes, allowing DNA to be introduced.

C. **Real-time PCR** – A real-time PCR machine amplifies and measures the production of amplicons in one step. It is a thermal cycler and fluorescent analyzer in one instrument and is usually computer-controlled. You do not have to load your product onto a gel to determine if it was made; the machine measures its production photometrically.

Techniques of Biotechnology and Innovations

The techniques used in modern biotechnology will be briefly highlighted in this section. A basic knowledge of their fundamental principles and applications is important for understanding the way biotechnology is used to benefit life. The techniques used in modern biotechnology can be divided into three major categories: genomics, proteomics, and metabolomics. Genomics is described as the study or use of genes and their functions. It involves techniques that investigate or make use of DNA. Proteomics is the study or use of the structure and function of proteins. It includes the various ways that proteins work individually and interact with each other inside cells. Metabolomics is the study or use of specific cellular processes carried out inside and outside of a cell. It includes the interaction of the cell with other cells (physionomics) and with environmental factors (environomics). Biotechnology techniques can also be classified into categories that analyze or apply the genomics, proteomics, and metabolomics. Analytical methods are used to analyze the function and structure of DNA, proteins, or metabolic pathways. These techniques today rely in laboratory instruments that detect the activity and chemical configuration of biological molecules. It was founded on the science of analytical chemistry, which is the analysis of chemical samples to gain an understanding of their chemical composition and structure. Application methods vary greatly and involve specific techniques for each category of biotechnology. Genomics applications require the modification of DNA. Procedures that control the functions or alter the structures of biological molecules are used in applications of proteomics and metabolomics. The manipulation of an organism's genomics, proteomics, and metabolomics is one of the fastest growing areas of biotechnology. New techniques are being developed every year for a variety of applications ranging from agriculture to pharmaceutical production.

ISOLATION OF NUCLEIC ACID (DNA)

Fundamentally all living organisms are either unicellular or multicellular, which can grow, reproduce, process the information, respond to stimuli and process the several other chemical /biological/physical process. These all processes are governed by genetic information stored in the genome of an organism. Either prokaryotic of eukaryotic organism's their genome is made of deoxyribonucleic acid (DNA). DNA is duplex of polynucleotide strands and it consisting of genes, promoter non-genic region and regulatory elements. To study this information and to characterize the genome of an organism, it is essential to isolate the genome (DNA) from their respective organism.

The total genetic materials vary from organism to organism and within the genera also. The DNA isolation methods more or less remain the same irrespective of size of the genome. However, the predominant factors like age of the sample and tissue type influence the isolation efficiency. The matured sample yield poor DNA quality than tender sample because of accumulation of phenolics like compounds.

The DNA isolation can be done in three phases. 1) Cell lysis, 2) DNA extraction and 3) DNA precipitation. General protocol is mentioned in Fig. 1.

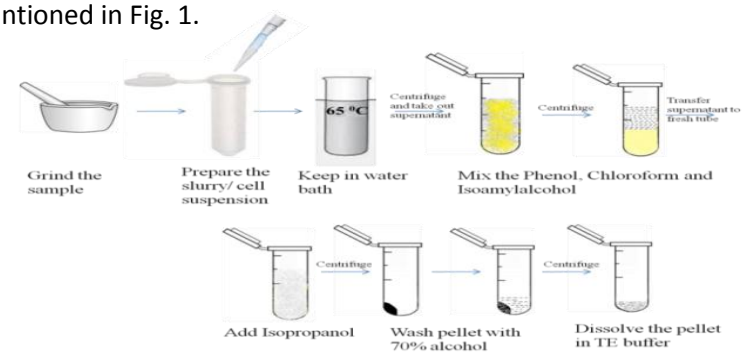


Figure 1: General procedure for plant genomic DNA isolation

Phase 1) Breaking the cell wall and cell lysis

The purpose of this step is to break-down the cell wall of the cells. For good quality of DNA the sample should be young and fresh. In case of plants the 4- 6 tender leaves (0.5-1 gm) are plucked and plugged into ice, for bacterial culture the log phased grown (0.4-0.6 Optical density at 600 nm) and young mycelia from fungal strain are used for DNA isolation.

Phase 2) DNA extraction

This step is performed to extract the DNA from prepared slurry suspension in the previous step. Once the cell is crushed the cellular materials such as protein/ carbohydrate is released into extraction buffer. These compounds negatively affect further enzymatic reaction like restriction enzymes digestion and polymerase chain reaction (PCR). Hence, the removal of these compounds is essential and performed by adding the chloroform and isoamyl alcohol (24:1) or phenol, chloroform and isoamyl alcohol (25:24:1).

Phase 3) DNA precipitation

The isopropanol will cause the DNA strands to condense and become visible and it is denser than the solution that they are in. The reaction is allowed to happen in chilled condition and later precipitated by centrifuging it.

DNA isolation from plant leaves

1. The leaf samples are crushed in mortar and pestle by use of liquid nitrogen and β-mercaptoethanol is mixed with 500 µl of CTAB buffer to prepare the slurry.
2. The slurry collected in microcentrifuge tubes is incubated at 65^{o}C for 30 min. During the process, the buffer components inactivate the DNase released from the cells.

3. The slurry is centrifuged at 10, 000 rpm for 10 min at 4^{o}C to remove the cell debris and the aqueous layer containing clarified cell suspension is transferred in to a fresh tube.
4. The Chloroform and Isoamyl alcohol in 24:1 ratio is added to collect aqueous layer from previous step in an equal amount.
5. The tubes are centrifuged at 10,000 rpm for 10 in at 4^{o}C and the aqueous layer is transferred to another tube without getting any contamination from middle and bottom layer.
6. An equal amount of pre chilled Isopropanol is mix with the collected aqueous layer of DNA solution from previous step.
7. It is mixed properly by inverting for couple of times and incubated for 30 min to overnight in deep freezer (-20^{o}C).
8. Later, it is centrifuged at 10, 000 rpm for 10 min at 4^{o}C and the collected pellet at the bottom of the tube is washed with 70% alcohol and it is spun for 2-5 min at 10, 000 rpm and the pellet is air dried.
9. The pellet is dissolved in 20-50 µl of TE buffer and stored in -20^{o}C till future use.

DNA isolation from *E. coli* (DH5α) bacteria

1. In bacterial sample, the log phase cells (15 mL) are harvested and dissolved in 10mM Tris-Cl (3 mL) and 100 mM NaCl (3 mL).
2. The suspension is centrifuged at 10, 000 rpm for 10 min at 4^{o}C and the pellet is re-dissolved in 2.5mL of TE buffer with 0.5 mL of 10mg/mL of lysozyme. The prepared cell suspensions are incubated at 37^{o}C for 30 min by gently agitating occasionally.
3. 25 µl of RNase (10 mg/mL) is added and incubated at 37^{o}C for 30 min.
4. 2.5 mL of SDS, which is prepared in 2 % TE buffer is added and incubated at 50^{o}C for 45 min

5. Protein is removed by adding the 50 µl of Proteinase-K (20 mg/mL) and reaction is allowed to proceed at 55^{o}C for 10 min.
6. The cleaved protein molecules are removed by mixing 6 mL of Phenol with previously collected material, and then centrifuged at 10,000 rpm for 10 min at 4^{o}C.
7. The aqueous phase is transferred to fresh tube and equal amount of phenol and chloroform in 1:1 ratio is added. The tube is centrifuged and the aqueous layer is collected. Chloroform and Isoamyl alcohol in ratio of 24:1 is added and the aqueous layer is collected from previous step.
8. The tubes are centrifuged at 10, 000 rpm for 10 min at 4^{o}C.
9. 1/10th volume of 3 M Sodium Acetate is added to the aqueous layer collected from previous step.
10. After incubating in ice for 20 min equal volume of Isopropanol is added and incubated at room temperature for 5 min.
11. The pellet is centrifuged and washed in 100 µl of 70% Alcohol.
12. The pellet is dried in 25-30 µl in $T_{10}E_1$ and stored in -20^{o}C for long term use.

DNA isolation from fungus (*Trichoderma*)

1. The fungal culture is incubated in potato dextrose broth (PDB) in 28^{o}C for 3-5 days.
2. The 3 gm of fungal mat is harvested from edge of the growing fungal culture, homogenized in mortar, and pestle in 4 mL of 2% SDS for 5 min.
3. 6 mL of lysis buffer is added to the above supernatant.
4. To this mixture RNase (10 mg/mL) is added and incubated for 30 min in 37^{o}C.

5. Equal amount of phenol, chloroform and isoamyl alcohol in 25:24:1 ratio is added. It is centrifuged at 10,000 rpm for 10 min at $4^{o}C$.
6. The supernatant is transferred to fresh tube and $1/10^{th}$ volume of 3M sodium acetate and equal volume of isopropanol is added, it is mixed gently by inversion and incubated on ice for 30 min.
7. It is centrifuged at 10,000 rpm for 10 min at $4^{o}C$ and the pellet is washed with 70% alcohol and the pellet is air dried.
8. The dried pellet is dissolved in 20-50 µl of TE buffer and stored in $-20^{o}C$ for future use.

DNA quality and quantity estimation

The purity and quantity of DNA is measured by spectro photometer or its purity is visualized on agarose gel. The DNA absorb the light wave length at 260 nm, the amount of absorbance is directly proportional to the amount of DNA present. The DNA sample is placed in quartz cuvette in order to read the OD. Based on absorbance coefficient the concentration of DNA is calculated. At A260 if the absorbance is 1 than concentration will be 50 µg/ mL and accordingly in other sample is calculated. For example, OD of DNA is 0.25 absorbance at A260 than concentration of DNA is 12.5 µg/ mL (0.25 × 50). At the mean time absorbance at 280 will give quantity of RNA in the sample. Hence, the A260/A280 gives the purity of DNA. If the ratio is 1.8 it indicate pure DNA, >1.8 shows presence of RNA and <1.8 shows protein contamination. It is visualized on horizontal agarose gel electrophoresis. In general, 1% of agarose is sufficient to visualize the genomic DNA. The gel is prepared by melting the 1 gm of agarose in 100 mL of TAE/TBE buffer. As the gel cools down the 3 µl of Ethidium Bromide (1 mg/mL) is dissolved and casted in the casting tray. The gel tray is immersed in buffer tank and ready for the DNA sample loading. The 2-3 µl of DNA is

mixed with DNA gel loading dye (BPB/Orange dye) and load into wells formed in the gel tray. Electrophoresis of gel for 45min-60min and visualize under the UV light. The presence of intact high molecular weight band indicates the good quality of DNA and smeared band indicates degraded DNA. The good quality DNA further used for different purpose such as molecular mapping, isolation of gene and blotting.

PLASMID ISOLATION

The method of plasmid isolation by alkaline lysis method was originally developed by Brinboim and Doly (1979). To isolate the good quality and quantity of plasmid the population of bacteria is important and growth phase. The bacterial cells are grown for log phase (0.4-0.6 O.D) at 600 nm in an appropriate medium and in appropriate conditions. Luria broth is congenial for *E. coli* DH5α cells, and 2xTY for *E. coli* TG1 cell. The bacterial culture is grown at 37^{o}C in shaking incubator at 200 rpm for overnight. These bacterial cells are lysed with alkaline Solution-I (Glucose 50 mM, Tris-Hcl 25 mM and 10 mM EDTA pH 8.0) and freely released protein and genomic DNA is denatured by adding alkaline Solution-II [0.2 N NaOH, 1% sodium dodecyl sulfate (SDS)]. NaOH, increases the pH and facilitates for denaturation, in contrast to the pH is decreased by adding the alkaline Solution-III (5 M potassium acetate- 60 mL, Glacial Acetic acid- 11.5 mL and sterile water-28.5 mL). This solution renature the plasmid alone but the genomic DNA will be linear and because of its larger size it unable to renature in short time of incubation. Hence, the denatured DNA precipitates and removed from the solution. The plasmid from the solution is precipitated by isopropanol and dissolved in sterile water/ $T_{10}E_1$ buffer. The detailed protocol discussed below.

Inoculate the individual colonies in 1 mL of LB with appropriate antibiotic selection pressure and incubate the cultures at 37^{o}C in 200

rpm. The overnight grown cultures are centrifuged at 10,000 rpm for 1 min at 4^{o}C. The supernatant are removed and the pellet is washed with 250 µl of STET. The suspended cells in STET buffer are centrifuged at 10,000 rpm for 1 min at 4^{o}C. Collect the pellet and dissolve in 100 µl of ice-cold alkaline lysis solution-I by vigorous vortexing. Later, add the 200 µl of freshly prepared alkaline lysis solution-II to each tube and the contents are mixed by inverting the tubes for 4 to 5 times and keep it in ice for 5 min. To this suspension, add the ice cold 150 µl of alkaline lysis solution-III and again mixed thoroughly by gently inverting the tubes for 4-5 times. Keep the tubes in standing position for 3 min. and spin for 10,000 rpm for 10 min at 4^{o}C. The supernatant is needed to transfer to fresh tube and add 2 µl RNase (10 mg/mL) and incubate for 30 min. at 37^{o}C. Later, add the Phenol, Chloroform and Isoamyl alcohol mixture in the ratio (25:24:1) in an equal amount to remove the protein and other moieties from the suspension. Mix the suspension for couple of times by inverting the tubes. Further, centrifuge at 10,000 rpm for 10 min at 4^{o}C. The aqueous layer is transferred to a fresh 1.5 mL pre chilled centrifuge tube and two volumes of isopropanol is added. The contents are mixed by inverting the tubes for 4 to 5 times and incubate the tubes in standing position for 5 min at room temperature. During the incubation the plasmid will precipitate and it is collected in pellet form after centrifuging, the suspension at 10,000 rpm for 10 min at 4^{o}C. Then discard the supernatant and the wash the pellet with 70% ethanol and spin for 1 min at 10,000 rpm to recover the plasmid. Discard the supernatantand collect the pellet at the bottom of the tubes and after its completely drying, dissolve in 25-30 µl of $T_{10}E_1$ (pH 7.4) (Sambrook and Russel, 2001).

RNA ISOLATION AND mRNA PURIFICATION

Preparation of tips, tubes and glasswares for RNA isolation

All the materials such as micropestle, tubes, tips, glass bottles, forceps, culture tubes, tips boxes and tube boxes used for RNA isolation are treated with diethylpyrocarbonate (DEPC) water to make them free from RNase. One mL of DEPC was mixed with 999 mL of autoclaved Nanopure water (0.1% of DEPC water) and incubates the bottles on rotor shaker at room temperature overnight for uniform dispersion of DEPC in water. The water was decanted and materials were dried in the hot air oven. These RNase free materials were autoclaved at 121^{o}C, 15 lbs for 15 min and dried.

Cell lysis and phase separation

The *Trichoderma* mycelia were crushed with liquid nitrogen in RNase free 2 mL micro centrifuge tube. The frozen materials were immediately subjected for cell lysis by use of TRIzol reagent (1 mL/100 mg of sample). After addition of TRIzol reagent the tubes were inverted for couple of times and incubated for 20 min at room temperature. Then 200 µl of chloroform was added to sample, mixed well and incubated for 5 min at room temperature. The sample was centrifuged at 10,000 rpm for 10 min at 4^{o}C. Within the tube TRIzol reagent separated the DNA, RNA and protein into three separate layers. The bottom layer considered of protein, middle layer consisted of DNA and top aqueous layer will contain RNA. Carefully upper layer was transferred to fresh pre chilled 1.5 mL RNase free tube without getting contamination from other layers.

RNA precipitation

The tubes were placed back in to ice as soon as aqueous layer was transferred. An equal amount (600 µl) of chilled Isoproponol was added to each tube and mixed gently by inverting them for couple of times. The samples were incubated at -80^{o}C for 2 h. Later it was

centrifuged at 12,000 rpm for 30 min at 4^{o}C. A very small gel like white pellet will be visible at the bottom side of the tube.

RNA wash

One mL of 75% alcohol was added to each tube and gently mixed and centrifuged at 10,000 rpm for 5 min at 4^{o}C. The liquid was decanted and allowed the tubes for complete drying of alcohol on tissue paper.

RNA solubilisation

The RNA pellet was air dried for 5-10 min and care was taken not to completely dry the pellet. Add 30-40 µl of nuclease free water to tube and the pellet was dissolved by repeated pipetting with a micropipette at 55^{o}C for 10-15 minutes. Later the RNA sample was stored at -80^{o}C until next use.

Restriction digestion

Restriction enzymes are nucleases which are able to cleave the DNA at specific sites. These are predominantly employed by microbes to restrict the multiplication of foreign DNA/ virus particles. Among the three different types of restriction endonucleases, the restriction endonuclease (RE) type II is predominantly using for most of the cloning experiments because of their site specific cleavage at recognition site and they require only Mg^{2+} as cofactor for their activity. The 1 U of enzyme is able to restrict the 1 µg of DNA in 60 min. Hence, depending on concentration of DNA the enzyme concentration to be use is various. Most of the RE works better at 37^{o}C. However, there are few exceptions and some shows star activity at a prolong incubation or at excess enzyme concentration such as *Bam*HI. Therefore, the enzyme concentration and incubation timing are most important in restriction digestion. The restriction digestion performed in microfuge tube containing DNA, RE, buffer and water. The buffer system is specific to RE and some RE requires the 1 X concentration of BSA as a cofactor. The restriction is performing with two different enzymes than it can be performed by

either double digest restriction or sequential digestion it is based on buffer system compatibility. The standard restriction performed for 60-90 min and depending up on enzyme property in activation will be performed like heat inactivation (65^{o}C for 15 min) or addition of 1/10th volume of 0.5 M EDTA but for some enzymes inactivation is not required. Based on the specificity and co-factor required Restriction enzymes are classified into three major classes, described below.

Types	Co-factors	Cleavage site	Enzymes
Type I	ATP, AdoMet, Mg2+	Their recognition sites and cleavage sites are different and cleavage sites are not fixed.	*EcoB, EcoK*
Type II	Mg2+	They cleave DNAs at the recognition sequence or at a defined distance from the recognition site.	*EcoR I, BamH I*
Type III	ATP, Mg2+	They cleave DNAs at fixed sites, though their recognition sites and cleavage sites are different.	*EcoP I, Hinf III*

Among these restriction enzymes, type II enzymes are generally used in genetic engineering experiments and widely available in the market. Up on restriction type II class enzymes, few produce the blunt end and few produces the sticky/cohesive ends (Fig. 2). Therefore, the first thing in restriction digestion is need to decide, which enzyme has to use and their compatible buffer for their action. Single molecule of DNA can be digest with one or more than one restriction enzymes (generally two) if the compatible buffer available that is called double digest, if the compatible buffer is not available than serial digestion can be used. The detailed protocol of the restriction digestion is described below.

1. **Select restriction enzymes to digest the target plasmid**
   To decide the type of restriction to use, first we need to know the sequence of the DNA molecule. Based on that, type of restriction enzymes can be decided.
2. **Determine an appropriate reaction buffer**
   Each enzyme works best at optimum pH, this pH and the co-factor required for the enzyme action is provides by buffer. Usually, it will be supplied by restriction enzyme suppliers.
3. **Template concentration**
   1 µg of DNA can be cleaved by one unit (U) of an enzyme in 1 h. Therefore, accordingly amount of enzyme required for amount of template has to decide. The amount of DNA that needs to be cut depends on the application. Diagnostic digests typically involve ~500ng of DNA, while molecular cloning often requires 1-3µg of DNA. The total reaction volume usually varies from 10-50µL depending on application and is largely determined by the volume of DNA to be cut. The template should be free of restriction inhibition compounds such as salt, DMSO and phenols.
4. **Sterile water**
   Water need to be pure and free of nucleases, in general the molecular biology grade water can be used. The amount of water has to add varies and depend on volume of other compounds addition and final reaction mixture.

Typical standard reaction (20 µl) consist of following reagents

- DNA – 1 µg
- Restriction Enzyme(s) – 1 U
- Buffer – 1X
- BSA / triton-X-100 (if recommended by manufacturer) – 1%
- Distilled water – to make up to total volume

Mix the above mixture gently and incubate for 1 - 2 hr at desired temperature (varies from enzymes to enzymes, usually 37^{o}C). After the reaction, some of the enzymes need inactivation either by chemical or heat (65^{o}C for 15-20min).

Need to take care:

- While keeping the reaction, do not add the enzyme directly on the buffer, it may cause irreversible damage to enzyme. Hence, first add water than buffer and BSA (if it is recommended). The enzyme needs to add last.
- Restriction enzymes must be placed in an ice bucket immediately after removal from the -20°C freezer because heat can cause the enzymes to denature and lose their function.
- The amount of restriction enzyme want to use for a given digestion will depend on the amount of DNA want to cut.
- If your enzyme did not cut, check to make sure that it isn't methylation sensitive.
- Sometimes enzymes cut sequences which are similar, but not identical, to their recognition sites. This is due to "Star Activity" and can happen for a variety of reasons like excess enzyme usage, excess time of incubation and high glycerol concentration.
- If the reaction is kept for more number of sample than it is better to prepare the master mix (all the component except template).

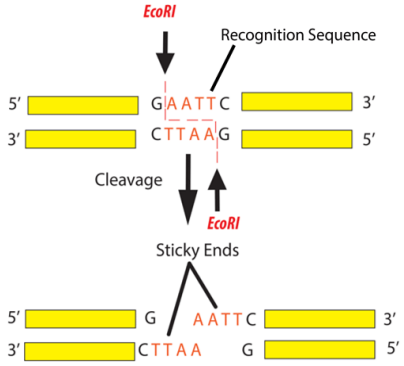


Figure 2: Recognition sequence of Enzymes

GENE ISOLATION

1) **Gene isolation by complementation**
2) **Positional cloning**
3) **Gene isolation by sequence homology**
4) **Gene isolation by functional genomics approach**

1) Gene isolation by complementation

The technique requires the genomic DNA/ cDNA library of functional alleles and host strain with deficit in trait/gene of interest. The success of technique depends on coverage of genome in the library. Therefore, library construction should be taken at most care and constructed in 2 different vectors to avoid the systemic exclusion. The each recombinant DNA vector from the library is transferred to host strain which does not carry functional allele of the gene of interest or may be recessive allele/mutant version of gene of interest. Screen/select the transformants by providing the conditions such as the gene of interest (GOI) should be present for transformants survival. Those does not able to complement the mutant host strain will die and those able to complement the mutant host strain will survive. For example, the

isolation of histidine gene from yeast, the genomic library transferred to mutant host and a cell/s that survives on minimal medium lacking histidine (His^-) that presumed to contain histidine gene of yeast. This particular clone can be sequenced and the complete gene can be obtained. The cloning by complementation does have some pit falls such as desired sequence may contain a region or encode a product that is toxic to the library for host organism and multiple suppression. The presence of multiple copies of gene (high copy number of plasmid) cause abundant amount of the product of one gene, it may compensate for loss of another different gene product. It may overcome the mutant/loss of gene function and cells may survive. It leads to wrong identification. Therefore to overcome this problem the copy number of vector need to be kept in mind.

2) Positional cloning

The positional cloning is the process of identification of physical position/proximity site of a cloned DNA fragment in the genome of an organism to isolate the gene of interest. The technique predominantly use chromosome walking process, where a cloned DNA fragment is a starting point from which the walking proceed step wise (clone by clone) fashion towards the gene of interest. During the walking a fragment of clone will be used as a probe to screen next overlapping fragment. The accuracy of determination of consequent fragment increased by use of markers like RFLP and STS and it will avoid wrong overlapping fragment identification. Therefore, physical marker is essentially required to this technique. To arrive at the site of gene of interest is calculated by measuring the distant / cross over frequency between the probe and gene of interest (GOI) trait. As cross over frequency reduces, it indicates the walking is towards the gene.

3) Cloning by sequence homology

Homology search based technique is simplest form of gene isolation. From the evolutionary point the genes are conserved between species or within the species hence the orthologous genes are used to isolate the genes from the desired organism. The primer or probe is designed from orthologous gene to amplify/ hybridize the GOI in the genome / genomic library respectively. Apart, if protein sequence is available instead of DNA than degenerative primers are designed to amplify the GOI.

4) Cloning by functional approaches

It involves the disrupting the GOI by a DNA fragment (may be T-DNA/TE) it results in loss of function. Screen for the desired mutant phenotype exhibiting individual and trace back to where the T-DNA / TE is inserted (the process is called gene tagging) and based on this the flanking sequence of TE can be identified and complete gene will be isolated.

MOLECULAR MAPPING

Molecular mapping is a technique used to locate the markers in the genome. The marker is feature of nucleotide in the genome. The technique facilitates to hasten the transfer of desirable genes between varieties and to identify the novel genes from wild species in to the cultivated crops. These molecular markers are method of analysis such as hybridization based molecular marker and PCR based molecular marker.

1) Hybridization based molecular marker

The technique is based on principle of hybridization between two complementary DNA molecules. Restriction fragment length polymerase (RFLP) is one, widely used hybridization based molecular marker. The genomic DNA is fragmented into small fragments with

restriction enzyme/s. The restricted fragment blotted on nylon/ nitrocellulose membrane from DNA resolved on agarose gel. The immobilized DNA fragments were hybridized with labeled probe (s). The labeled probed that hybridizes with one/more of the blotted DNA sample, thus revealing a unique blotting pattern characteristic to a specific genotype at specific loci. The amplified RFLP product resolved in agarose/polyacrylamide gel electrophoresis and based on banding pattern presence/ absence of a characters or similarity between genotypes can be studied.

2) PCR based marker

PCR is a molecular biology technique for enzymatically replicating small quantity of DNA without using a living organism. Here an oligonucleotide fragment of 20-24 nucleotides (primer) designed. The primer can anneal and amplify a specific DNA fragment (site directed PCR markers) such as EST, CAPS, SSR, SCAR, STSs or primers can amplify the random region in the genome (arbitrary /semi arbitrary primed based PCR) such as AP-PCR, DAF, RAPD, AFLP and ISSR.

GENOMIC DNA LIBRARY CONSTRUCTION

The collections of recombinant DNA molecule carrying the insert of genomic DNA fragment of organism, the sum of total DNA insert of this collection represents the entire genome of the respective organism. For the library construction entire genomic DNA of an organism subjected for physical sharing or enzymes based digestion with the intention to produce random fragments. The partial digestion of genomic DNA is performed with frequent cutter restriction endonuclease, mainly 4 bp cutters. It is assumed that the recognition sequence of restriction enzyme arranged randomly hence chance of occurrence of 4 bp restriction enzymes site is once in 4^4= 256 bp. The cloning of the small fragment leads to achieve high rate cloning efficiency with such enzyme

partial digestion performed either by giving half of the normal enzyme quantity or reducing the time of incubation to half of the normal timing. The produced fragment are cloned into the suitable vectors (plasmid/ phage based vectors) depending upon the size of the fragment produced. In general, lambda based or cosmid vectors are used to clone the insert size of 23-25 Kb and 50 Kb respectively. The recombinant molecules are transferred in to an appropriate host. The transformant number may vary and depend on size of the fragment produced from intact genome. Hence, size of the library directly proportional to number of fragment produced. Further, it depends on type of restriction enzyme used. For example, genome size of an organism is 1,00,000 bp and 4 bp cuter is using for genome library construction. Then, number of clones produced is directly proportional to genome size and inversely to size of fragment produced from restriction enzyme. For 4 bp cutter total 340 cones can be produced from 1, 00,000 bp genomic DNA. Apart, the cloning efficiency and probability of insert representation in the library cost the library size.

TRANSFORMATION TECHNIQUES

The uptake of foreign DNA by cell is called transformation. This technique incorporates a new DNA fragment into a host cell. Therefore, incorporated cell will get extra trait/ improved its one of the trait. It can be achieved by two methods 1) *Agrobacterium* medium transformation 2) Direct gene transfer. Both the techniques aim at stable integration of foreign DNA molecule into the plant genome. Hence, the incorporated gene will inherent from generation to generation.

***Agrobacterium* mediated transformation**

The binary / co-integrated vector carrying the GOI transferred into the plants by *Agrobacterium* in two ways 1) co-culturing 2) In plant transformation.

1) Co-culturing

The *Agrobacterium* carrying GOI carrying T-DNA and *vir* plasmid is cultured with plant cell/s (explants). The explants are the part from which has potentiality to regenerate into whole plant. Explants may be cotyledon leaves, thin cell layers, protoplast, tissue slice, leaf disc and section of roots, shoots or floral tissues. Before the co-culturing with *Agrobacterium* the explants are surface sterilized with 0.1% mercury chloride for 1 min (varies with explants and type of crop). The traces of mercury chloride washed off by sterilized water for 3-4 times. The dried explants incubated with log phase grown desired gene carrying *Agrobacterium* culture. The culture incubated for 2 days under dark condition during the acetosyringone released from tissue (dicot) or artificially added (monocot) induce the *vir* genes expression. The series of *vir* protein act on T-DNA containing GOI and kanamycin resistant gene (antibiotic resistant gene) to transfer the gene/s into cultured explants genome. Later, cultured explants are washed in sterile water for couple of time to remove the *Agrobacterium* from explants. The washed explants are dried in sterile blotting sheet and later kept in regeneration medium containing antibiotic selection pressure such as carbenicillin to restrict the *Agrobacterium* growth and kanamycin to screen the transformed cell. In about 3-4 weeks the shoot will regenerate and it is transferred to rooting medium. In 15 days roots are produced and these plants are subjected for hardening. The two level of hardening carried primary hardening and secondary hardening. In primary hardening the tissue culture planting are transplanted into peat and kept under humid

condition with temperature of 25±1 after 2 weeks these plants are transferred to soil / green house.

2) In plant transformation

This technique circumvents the tissue culture practice and direction the sterilized plant tissue will be soaked *Agrobacterium* culture. Tissue may be seed, flower and leaf, tissue is immersed in a fresh culture of *Agrobacterium*, for an efficient transformation the vacuum is created, and it will facilitate the *Agrobacterium* to enter the cell. The plants are grown from these seeds and progeny obtained from this is believed to be containing the GOI in its genome.

DIRECT GENE TRANSFORMATION

Introduction of DNA molecule into the cells by means of chemical and physical methods is called as direct gene transformation. The naked DNA molecule cloned or un-cloned is directly used for delivering into the cells. The varies methods of direct gene transformation are 1) chemical 2) biolistic transformation 3) electrophoresis 4) lipoprotein 5) microinjection 6) laser induced and 7) fiber mediated gene transfer methods are in use.

Probe preparation

The probe preparation includes two major aspects 1) template preparation and 2) labeling the probe. The template may be of double stranded DNA (dsDNA) probe, single stranded (ssDNA) probe, RNA probe (riboprobes) and synthetic oligonucleotides. These probes are labeled with radioactive molecule (^{32}P, ^{35}S and ^{3}H or non-radioactive molecule (biotin, digoxigenin and fluorescent dye).

LABELING METHODS

Nick translation:

The technique provides labeling of DNA by radioactive or non-radioactive labeling molecules. The reaction involves two enzyme action where, one of it produces nick in single/double strand and another enzyme extend the DNA strand from free 3' -OH group of nick end and it extend and replace the existing nucleotide from 5' side of the nicked DNA fragment. The pancreatic DNase I makes a nick and *Escherichia coli* DNA polymerase I (Klenow fragment) to extend and labeled nucleotide into DNA strand. The technique produces the dsDNA probe. The reaction prolong for 1-2hrs depending up on type of enzymes and their efficiency. The reaction mixture consist of 1U of DNase I, 5-15U of DNA polymerase I, 1x concentration compatible buffer, nuclease free water, 0.25 – 1 µg of template and dNTP mixture. For radioactive labeling the three dNTPs (dATP, dTTP and dGTP) are mixed in equal concentration and [α -32P] dCTP (3000 Ci/mmol, 10 µCi/µL) are mixed and for non-radioactive labeling the [α -32P] dCTP is substituted with fluorescent molecule labeled nucleotide such as biotin-11-dUTP, fluorescein-12-dUTP, DIG-dUTP or aminoallyl-dUTP molecule is used. During the reaction along with non-labeled nucleotides (1 mM dATP, 1 mM dCTP, 1 mM dGTP, 0.65 mM dTTP) mixed with 0.35 mM DIG-11-dUTP (molar ratio of 1:3-1:5). Incubate the reaction mixture at 15^{o}C for 2 hrs and later it is stopped by addition of 1µL of 0.5 M EDTA, pH 8.0. These labeled molecules are extracted from reaction mix by DNA precipitation method.

Primer extension

The reaction is a normal PCR method, where it includes the one of the nucleotide is labeled and polymerase enzyme is Klenow fragment. The reaction involves the denaturation of double strand DNA into single strand DNA and the random primer (hexamer) of 6-10 base long in

allowed for annealing. The one of the labeled dNTP such as biotin labeled dCTP or DIG labeled dUTP incorporated in to synthesizing DNA strand. The incorporation take place by complementary base pairing method hence, the incorporation of labeled molecule take place all along the DNA strand.

End labeling

The nucleic acid's either 5' or 3' ends of the strands are labeled. The 3' end labeling involves the tailing reaction by terminal transferase, which is template independent DNA polymerase. The normal PCR runs with unlabeled nucleotide along with labeled fluorescent labeled biotin/ DIG-11-dCTP molecule. The terminal transferase incorporates this labeled molecule at the end of the 3' end of the oligonucleotide. Similarly the 5' end can be labeled with radioactive/ non-radioactive molecules. The 5' end labeling in a two-step synthesis with first an amino linker residue on the 5' end of the oligonucleotide, and then after purification, a digoxigenin-N-hydroxy-succinimide ester is covalently linked to the free 5'-amino residue. Apart from this, it can also be labeled with radioisotopes by transferring the γ-^{32}P from [γ-^{32}P] ATP to the 5' end using the enzyme bacteriophage T4 polynucleotide kinase. Before addition the template strand is dephosphorylated by calf intestinal alkaline phosphatase and a labeled phosphate molecule incorporated at this site by polynucleotide kinase.

Riboprobe

This method includes labeling of RNA molecule by either radioactive or non-radioactive molecules. The target DNA (cDNA) fragment need to be labeled is cloned in to a bacterial vector system and expressed under the viral promoters like T7 or SP6 promoters. This molecule is expressed under *in vitro* condition by use of viral RNA

polymerase specific to target promoter. During transcription process the labeled molecules incorporated in to RNA molecule. For SP6 and T7 promoters their respective RNA Polymerases are incubated with 40mM Tris-HCl (pH 7.9), 10mM NaCl, 6mM MgCl2, 10mM DTT, 2mM spermidine, 0.05% Tween®-20, 0.5mM each of dATP, dGTP, dCTP and dUTP, 0.5µCi [3H] dCTP and 1-2 µg of the linearized template DNA/ vector molecule. The reaction is allowed to take place at 37^{o}C for 1hr.

SOUTHERN BLOTTING TECHNIQUE

The technique is useful in confirming the cloned fragment in the plasmid/ genomic DNA of corresponding organism. The template genomic DNA molecule fragment with a defined size by restriction enzymes and immobilized on nitrocellulose membrane. The labeled probe is allowed to hybridize with complementary fragment which is immobilized on the membrane. The technique includes 4-5 steps viz., isolation of DNA, blotting technique, DNA/DNA hybridization, washing and visualization (Fig. 3).

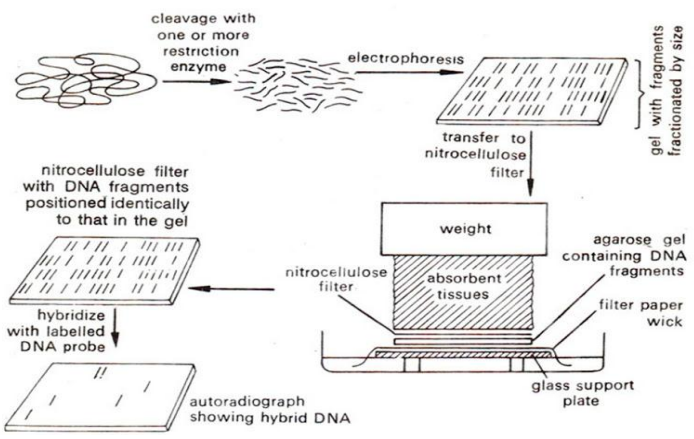

cleavage with one or more restriction enzyme
electrophoresis
gel with fragments fractionated by size
transfer to nitrocellulose filter
nitrocellulose filter with DNA fragments positioned identically to that in the gel
weight
agarose gel containing DNA fragments
nitrocellulose filter
absorbent tissues
filter paper wick
hybridize with labelled DNA probe
glass support plate
autoradiograph showing hybrid DNA

Figure 3: Steps in Southern Blotting

The DNA is treated with one or more restriction enzymes and fragments are separated on the agarose gel. Before transferring to nylon/nitrocellulose membrane the double strands fragmented DNA is denatured by alkaline treatment in order to facilitate the immobilization and proper hybridization with the probe. If the nitrocellulose membrane is being used for blotting than after alkaline treatment soak the gel in Tris salt buffer, because under highalkaline condition (pH 9.0) the DNA fragments will not bind to nitrocellulose membrane. The prepared gel placed on top of the buffer saturated filter paper, then thin layer of nitrocellulose paper on top of the gel and finally some dry filter paper kept over the nitrocellulose membrane. Some wait will be kept on the top to facilitate good capillary action. The entire assembly kept in buffered solution but only lower filter paper should touches to the buffer, it will build the capillary action and from the top dry paper receive the buffer. During capillary movement, the DNA transferred from gel to membrane, in the same fashion as they are present in the gel. This transfer buffer varies with type of membrane used, for nitrocellulose membrane the high-salt transfer buffer (20 X SSC, which comprises 3.0 mol/L NaCl and 0.3 mol/L sodium citrate) and for nylon membrane alkaline transfer buffer (0.4 mol/L NaOH) is used. The transfer process takes for 18 h in case of SSC buffer and 2 h in alkaline transfer buffer. After the transfer process gently rinse the membrane in 2X SSC buffer and air dry. At this step the DNA is loosely bound to nitrocellulose membrane hence, it is permanently immobilized by baking the membrane at 80^{o}C for 2 hr. The free space in membrane is blocked with 0.2% each of Ficoll, polyvinyl pyrrolidone and bovine serum albumin, the process is called as prehybridization, runs for 15 min to 3 h at 65^{o}C. Alternatively, this mixture can be included in the hybridization buffer. Later treat the membrane with hybridization buffer 2 X SSC containing 1% SDS and labelled (radioactive labelled/ fluorescent labelled) probe.

Keep it in agitating condition or in a tube that is constantly rotated so all parts of the membrane are exposed to the probe. During the hybridization chances of non-specific annealing, it can be reduced by using the buffer of high ionic concentration and incubating the hybridization process at/just below melting temperature of probe and template. After hybridization, the membrane is washed and subjected for detection procedure either by X-ray in case of radioactive labelled probe or naked visualization for fluorescent labelled probe.

***IN SITU* HYBRIDIZATION (ISH)**

The localization of a fragment of specific nucleic acid in portion or section of tissue by hybridization with cDNA or RNA molecule as a probe is called as *in situ* hybridization (ISH). In general, the probes are labeled either with radioactive molecule or fluorescent molecules and their hybridization to DNA indicates the presence or absence of target DNA molecule in genome and their hybridization to RNA indicates the gene expression information.

Steps involved in ISH

1. **Tissue preparation**
2. **Probe preparation**
3. **Hybridization and Visualization**

1. Tissue preparation

Harvest the target tissue and rinse with 1x phosphate buffer saline (PBS). The dried tissues is immersed in freshly prepared 4% paraformaldehyde-0.1M sodium phosphate buffer pH7.4 and incubate at 4^{o}C for 3hrs. The purpose of this step is not to completely fix the tissue but to harden and preserve the tissue. The complete fixation carried just before the hybridization which will fix all section equally. Later, tissues were immersed in 15% sucrose in 1x PBS **(**500mL sterile PBS, 75g "RNase

free" sucrose) for 3hrs to overnight at 4°C. Now, tissue is ready for sectioning and hence, tissues are embed in a matrix like parafilm in O.C.T blocks and used for sectioning. The tissues are sectioned on cryostat to 5-10 µm thickness, thaw-mounted onto coated slides and immediately flash frozen (-70°C).

2. Hybridization

The frozen tissues are mounted on slides and thawed for 5 min at 50oC this enhances the tissue retention on slide till the last step of the process. The dried slides are dipped in 4% paraformaldehyde **(**200mL 0.5M NaPO4, pH 7.4, 800mL DEPC H_2O, Heat to 70°C with stirring on hot plate in fume hood Add 40g Paraformaldehyde) for 10 min at 4oC, this ensures the equally fixing of tissue on slide and it permanent retention of tissue on slides. The paraplast coated to tissue is removed by rinsing the tissue with xylene solution for 10 min. the slides are dipped in 100% ethanol for 15 mins or until strikes of xylene removed from the slide. These tissues are hydrated with series concentration of ethanol 95%, 85%, 70%, 50%, 30%, 15% and H2O. Repeat the repeat the step until strikes go away and tissue is hydrated. Incubate the slide in 0.2N HCL for 20 min at room temperature to denature the proteins and nicks the DNA. Later the slides are incubating in 2x SSC (1x SSC- 150 mM NaCl, 15 mM sodium acetate, pH7.0) 70°C, for 15 min and wash the slides with 1x PBS for 2 min at room temperature. Before hybridization, to increase the probe penetration, the tissue membrane softens by treating the tissue with 1.0 µg/mL of proteinase K for 10 min at 37°C. Wash the slides with sterile water for couple of times and rinse with 1x PBS. Further protease action inhibited by incubating the slides in 2mg/mL glycine in 1X PBS for 2 min at room temperature. Further, the slides are rinsed with 1x PBS for 2 times for 2 min. The dried slides are kept in hybridization box covered with 4x SSC buffer and 50% formamide saturated wick. Initially the pre-hybridization carried out where, 100 µl of hybridization buffer (for

riboprobe hybridization-10 mM DTT, 0.3M NaCl, 20mM Tris pH 8.0, 5mM EDTA, 1x Denhardt, 10% Dextran sulphate, 50% formamide; for oligomer hybridization 10 mM DTT, 1x Denhardt, 5x SSC, 100 µg/mL of ssDNA, 100 µg/mL tRNA, 10% Dextran sulphate, 20% formamide) spotted on the tissue. The process goes for 3hrs at 42^{o}C and later hybridization begins by adding the 20 µl of probe solution. The hybridization is carried out at 55^{o}C for overnight incubation and next day, the slides are rinsed twice for 10 min with 2x SSC containing 10 mM β- mercaptoethanol and 1 mM EDTA at room temperature. Finally the tissue sections are dehydrated with gradient concentrated alcohol (15%, 30%, 50%, 70%, 85%, 95%, and 100%) containing 0.3M ammonium acetate and after bring it can be stored at room temperature. The hybridized probe can be visualized by epifluorescence microscopyfor fluorescent labelled probes and autoradiography used to observe the radioactive labelled probes.

3. Suppression subtractive hybridization (SSH)

SSH is a transcription profiling tool that aids in identifying the differentially expressed genes between two different (healthy and diseased) samples. The data explains type of genes that are differentially expressed in various biochemical pathways, regulatory pathways, signalling pathways, defense systems and adaptability. Apart from this, it co-relates the inter-relationship between identified genes, clustering of genes and understanding the cell behaviour at the molecular level. In SSH, mRNA population of one sample will be subtracted in mRNA population of other sample and differentially expressed mRNA copies are enriched. Usually the samples will be two counteracting samples. The sample in which genes are expected to express differentially is used as tester and its counteracting partner is driver. The SSH can be used in two different ways (i) forward (ii) reverse subtraction and both give different information such as upregulated and downregulated genes respectively. The technique explored in different biological system to identify and

study gene expression pattern under different conditions (Diatchenko *et al*., 1996).

Expression of luxury genes depends upon conditions, confronting timings and type of host confronting. Hence, it is essential to consider these parameters to discover differentially expressed genes in tissue/cell. The SSH performed by preparing the cDNA from special condition (tester) and it will be subtracted in cDNA of its respectivecontrol (driver) condition (non-confronting sample). In the beginning step, the tester cDNA is divided into two groups (pool A and pool B) and each group ligated with one type of adaptor. In first round of hybridization, the driver cDNA allowed to hybridize with both the pool of cDNA separately. During the process the common genes between the tester and driver samples will get hybridize. In next round of hybridization, the product of first hybridization (pool A and pool B) mixed without any denaturation, allowing the hybridization to take place between unhybridized cDNA of pool A with unhybridized cDNA of pool B. During the process, the differentially expressed genes will get hybridized and these molecules carry two different adaptors at their ends. When this sample is subject to PCR with the adaptors specific primers than, the differentially expressedgenes will (because they carry two different adaptors at their ends) amplify and enriched with differentially expressed genes only. As the technique facilitating to discover the new genes and understanding their role in spatial condition, it is gaining more importance in research field. The SSH involves series of steps starting from, cDNA synthesis to amplification of differentially expressed gene and briefly described below (Fig. 4).

Methodology

Methodology mentioned here is based on the work of Diatchenko *et al*. (1996) and SSH kit protocol of TAKARA, USA. The 2 µg

of poly A^+ RNA is required concentration for good quality and quantity preparation of cDNA. The primer specific to poly A^+ tail (10 µM) is mixed and incubate for 2 min at 70°C which will allow the denature the soiled RNA molecule and annealing of primer to their location. Cool the tubes on ice for 2 min, the template hybridized to primer is ready for cDNA synthesis. Later, incubate the reaction with reverse transcriptase (100 units/µl, 1 µl/reaction), dNTP mix (10 mM each), 1X buffer and sterile water for 1.5 h at 42°C. Terminate the reaction by placing the tubes on ice and immediately proceeded for second strand cDNA synthesis. The RNA:DNA hybrid is denatured and remove the RNA molecule by RNase H enzyme (2 h incubation) and second strand is synthesised by adding the T_4 DNA polymerase (6 U) per reaction. Incubate the tubes (in thermal cycler) at 16°C for 30 min, later the second strand synthesis was terminated by adding 4 µl of 20X EDTA/Glycogen mix and synthesised cDNA is precipitated and isolated by DNA purification method. The pellet is dissolve in sterile (43.5 µl) of water. The prepared double strand cDNA is digested with 1X of *Rsa* I enzyme as per the standard digestion procedure. It creates the blunt end double strands. Purify the digested short fragment by phenol: chloroform: iso amylalcohol method. The digested tester cDNA molecules is divide in to two group and each group attached with one type of adaptor with the help of T4 DNA ligase.

The adaptor ligated molecules are ready to hybridize with *Rsa*I digested control sample cDNA molecules. Total two rounds of hybridization is carried out *viz*., (1) in first round of hybridization the adaptor ligated tester cDNA molecules of both the pools are mixed with digested control cDNA molecule individually in separate tubes. (2) The product of the first round hybridized pool B sample is mixed with is mixed product of the first round hybridized pool A with extra control driver cDNA molecules. The non-luxury genes will get hybridize in first round of hybridization and luxury genes with two different adaptors will

hybridize in second round. During the first round of hybridization, the denaturation of sample is carried at 98^{o}C for 1.5 min and whereas, in second round of hybridization the denaturation will not be followed. For first round of hybridization 8 h of incubation at 68^{o}C and for second round of hybridization, the little longer time is required (10-12 h) at 68^{o}C. These molecules further amplified with two rounds of PCR with primer specific to adaptor sequence. Hence, the cDNA molecule those carry both the types of adaptor will get amplify exponentially and the product will be enriched with only differentially expressed genes.

The 1 µl of each subtracted cDNA (needed to be diluted) of sample and their corresponding unsubtracted cDNA aliquot (this sample is collected before the subtraction) are subjected for the PCR with PCR primer 1. Initially, the tubes were incubated at 75^{o}C for 5 min in a thermal cycler to extend the adaptors.In continuation of this incubation step, PCR program will commence (94^{o}C for 20 sec, 94^{o}C for 10 sec, 66^{o}C for 30 sec and 72^{o}C for 1.5 min. Dilute this product in 1:10 ratio and use 1 µl for second round of PCR. The nested primer used in second round of PCR it increase the specificity and only differentially expressed genes will exponentially amplify, the PCR program is 94^{o}C for 10 sec, 68^{o}C for 30 sec and 72^{o}C for 1.5 min. The amplified product is stored at -20°C and used for sequencing.

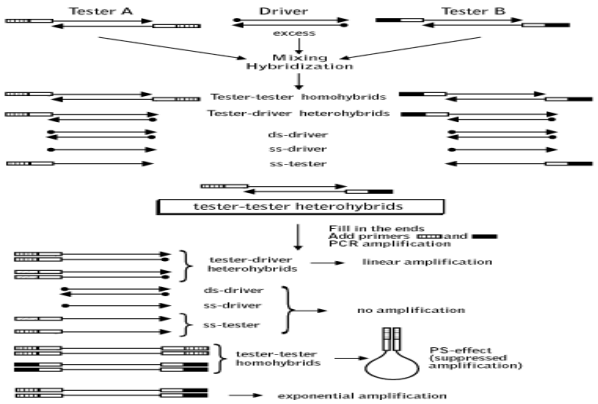


Figure 4. The graphical representation of SSH technique.

POLYMERASE CHAIN REACTION

Polymerase chain reaction is better known as PCR, the process is of chain reaction which amplifies the small quantity of starting material to larger number. The technique used to amplify the DNA fragment to simply create multiple copies of a portion of DNA or to compare two different samples of DNA to know which is the more abundant or to find out the similarity between the two/ more than two genotypes. The reaction is performed by DNA polymerase enzyme where it synthesis the complementary DNA strand by extending the DNA strand from free –OH molecule at the starting point. The free –OH molecule provides by a small piece of DNA strand, called as primer. The primer is an oligonucleotide about 10-25 bp in size and binds to its complementary DNA strand and facilitate to amplify it to large number. The process involves the series of steps viz., denaturation of DNA template, primer annealing and extension of DNA strand from free 3' –OH group of primer. These three processes are repeated for few cycles so that the copy number of the template increases. The three steps operate at their

own temperature like for denaturation of DNA requires the 94^{o}C to break the hydrogen bonds between the complementary strands and in contrast to it facilitating the condition to form the hydrogen bonds between the primer (~50^{o}C) and template. The annealed primer will be extended by polymerase enzyme (~72^{o}C). The polymerase enzyme produces the complementary strand by incorporating the complement nucleotide of the template to the existing free 3'-OH of the primer. Hence, series of amplicons will produce in a short time by repeating the cycle of temperatures. The first amplicon produced in the third cycle from than onwards its copy number doubles as the cycle progress for example- in fourth cycle 2 copies, in fifth cycle 4 copies, in sixth cycle 8 copies and so on. At the n^{th} cycle the number of copy produced will be $2^{(n+1)}$, where "n" is number of cycles.

Components of PCR

1) DNA polymerase

As the temperature in the reaction rises to 94-95^{o}C, the components used in the reaction should retain their activity. Hence, *Taq* polymerase isolated from *Thermus aquaticus,* which will not lose the confirmation at the higher temperature. In general, the 1 unit (U) of DNA polymerase able to amplify the 1 Kb of amplicon based on that the quantity of enzyme varies, 0.5-2.5 U of enzyme used for 20-50 µl of reaction. The different type of polymerase enzymes are available commercially those may varies with respect to fidelity and efficiency of polymerization.

2) Primer

The nucleotide sequences in the primer decide the region to be amplified. The primers may be 10 to 25 nucleotides; size varies with the purpose of the amplification. The arbitrary primers usually 10 nucleotides in size and primers to gene specific will be 18-

25 nucleotides long. The annealing temperature of the primer to the target template depends on the nucleotide composition of the primers. In general, a set of primer is used to amplify a region, like forward and reverse primer. For better amplification, the difference between the annealing temperature of forward and reverse primers should not be exceed the $5^{o}C$. As the polymerase begins with free –OH group from 3' end of the. The primer's 3' end most important than 5' and it should be perfectly anneal to the template and it is better to design with the G/C end. The 5' end is freely available, this end can be used to attach for any modifications like attachment of restriction sites/ nucleotide labeling. While designing the primers these points should be considered. The 5-10 pmol concentration of primers is routinely used for normal 20-50 µl PCR. Even the excess primer concentration inhibits the amplification sequestering the buffer component (Mg^{+2}) as a result *Taq* polymerase faces the shortage of buffer component (Mg^{2+}) for amplification.

3) Deoxynucleoside triphosphates (dNTPs)

The building blocks of DNA i.e., dNTPs, are adenine, guanine, thymine and cytosine are arranged randomly with meaningful functions in the DNA. Hence, while amplifying the DNA under *in vitro* conditions the all four nucleotide added to PCR in an equimolar concentration. The 1 mM-2 mM are generally used for each reaction of PCR and it sufficient to amplify the 1 kB of DNA template. High concentrations of dNTPs (>4mM) are inhibitory, perhaps because of sequestering of Mg^{2+}.

4) Buffer

The pH of buffer between the 8.3- 8.8 is congenial to *Taq* polymerase to carry out the amplification. The enzyme supplied in 10 X concentration, the 1 X concentration is sufficient for amplification.

Now-a-days, divalent cation is added to the buffer the concentration of 1.5 mM of divalent cation (Mg^{2+}) is sufficient for the PCR.

PCR programme

The PCR starts with initial denaturation of template DNA with 94-95^{o}C for 5-10 min depending upon the G+C content of the sample. 5 min is preferentially used which is sufficient to denature the genomic DNA. Followed by this step, cycle of different temperature begins. The cycle (Fig. 5), begins with 94-95^{o}C for 45 sec, ~50^{o}C for 45 sec facilitate for primer annealing (varies, depends on primer G+C content) and extension step for amplification of single strand it is performed at 72^{o}C for 45 sec. this cycle will be repeated for 28-32 cycle for sufficient quantity of amplicons. As the number of cycles progressed, the amplification efficiency comes down due to exhausting of substrate molecule of PCR. Hence, it is better to do with optimum number of cycles (28-32).

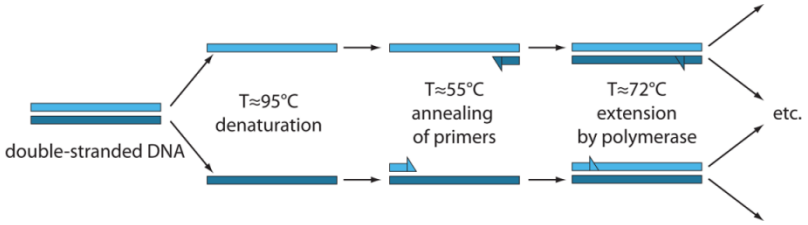


Figure 5: Steps of Polymerase Chain Reaction

qRT-PCR

Routinely the PCR products will be analysed in a separate procedure, which performed after the reaction completion. This type of analysis is called as end point analysis but this method is most suitable to present the presence or absence of a product on horizontal gel electrophoresis. The accurate quantification by this method possess

several disadvantages inaccurate quantification initial template copy number because as reaction progress to (30-40 cycle) the availability of polymerase substrate become limited for polymerase reaction and accumulation of inhibitors, hence the amplification will not progress in an exponential rate and amplification reaches plateau phase (Fig. 6). Quantifying after this phase yields poor results therefore it is essential to quantify cDNA template at an exponential phase, achieved by monitoring the amplification rate by use of a fluorescent dyes and measuring their signals by high resolution cameras. This procedure is known as real time quantification. SYBR green is most commonly using fluorescent dye that binds to only double stranded DNA and emits the fluorescence. In PCR as the double strand DNA synthesis progressed the binding of SYBR green dye to dsDNA increases and emits more fluorescence, which will be measured in real time later used for quantification of initial copy number fold change of a template by use of a mathematical formula (Ramakers *et al.,* 2003, Czechowski *et al*., 2004). Apart from reaction, the template integrity, concentration and primer concentration plays a major role in quantification.

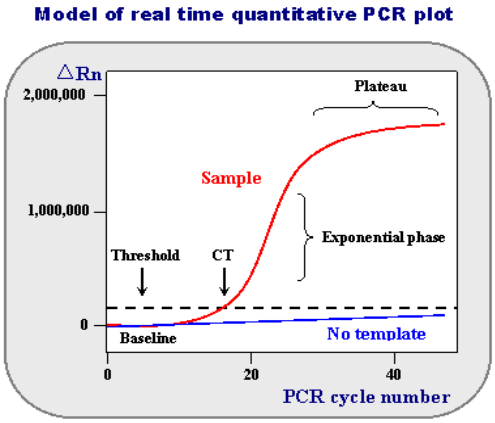


Figure 6: Schematic representation of qRT PCR graph

The exponential growth phase will be used for quantification of target gene concentration. Where, ΔRn is incremental in fluorescent signal at each time point, Ct is cycle threshold value. The ΔRn plotted against PCR cycle number to obtain PCR amplification curve.

Primer designing

The primers pairs can be designed by Primer3Plus software. A predicted melting temperature (*Tm*) of 60+2^{o}C, primer lengths of 20-24 nucleotides, guanine-cytosine (GC) contents of 45-55 % and PCR amplicon length of 100-200 bp is optimum for designing the primer pairs. The specificity of primer pairs can be further reconfirmed by searching homology in NCBI, BLAST search.

Template concentration standardization

In real-time quantitation of any gene expression, template concentration is one of the key factors to consider. The different concentration 10 ng, 1 ng, 0.1 ng, 0.01 ng and 0.001 ng of a gene need to optimize the amplification. Template concentration which gives the low threshold value (Ct) and high fluorescence value can be further used for remaining samples.

Primer concentration standardization

In real-time quantitation of any gene expression, primer concentration is one of the key factors.Primer concentrations of 1 pmol, 5 pmol, 10 pmol and 15 pmol are generally used to optimize the amplification. Primer concentration which gives the amplification with low Ct value and high fluorescence value is generally used for remaining genes.

qRT-PCR reaction

The master mix of different components of real-time PCR was prepared fresh to avoid handling errors. The reaction carried out in two technical replications and two biological replications. At each time, a minimum of 10 µl reaction mixture prepared containing desired concentration of cDNA template (1 µl), a pair of (5 pmol) gene specific primers (0.5 µl each), 5 µl of 2X SYBR green reagents and desired amount of sterile, nuclease free water. The PCR program followed was, initial denaturation temperature of 95^{o}C for 10 min, followed by denaturation at 95^{o}C for 15 sec, primer annealing temperature with annealing time 20 sec, extension at 72^{o}C for 20 sec. Run the reaction in thermal cycler instrument, which can be able to detect the fluorescent molecules (Eppendrof mastercycler *ep* realplex thermal cycler instrument).

Setting the baseline and threshold level

The accurate threshold level is measured with respect to point at which SYBR green fluorescence signals of amplification crosses the background signals. The background signals arise because of the changes in the reaction conditions, media and environment. During the PCR, at a cycle the signals will be high and cross the threshold line that particular cycle number will be measured by Eppendrof mastercycler *ep* realplex signal detection software and denoted as Ct value. The default value will be 3-15 cycles.

Relative gene quantification analysis

Relative quantification determines change in steady state mRNA levels of gene/s in a sample or multiple samples and express it relate to the level of internal control RNA. The most of the cases internal control will be of housekeeping genes which know to be express uniformly in cell even under different conditions. Relative quantification based on

expression level of target gene and internal control gene. Hence, preparation of calibration curve is not essential for this strategy. The various mathematical formulas are available to calculate the fold change in gene expression, mainly they are depends on (i) without PCR efficiency correction (Pfaffl *et al*., 2001) (ii) with PCR efficiency correction (Tichopad *et al*., 2003).

(1) Without efficiency correction:

$$\text{Ratio} = 2^{(-\Delta Ct\ \text{sample} - \Delta Ct\ \text{control})}$$

(2) With kinetic PCR correction efficiency:

$$\text{Ratio} = \frac{(E\ \text{target})^{\Delta Ct\ \text{target (control-sample)}}}{(E\ \text{ref})^{\Delta Ct\ \text{ref (control-sample)}}}$$

or

$$\text{Ratio} = \frac{(E\ \text{ref})^{Ct\ \text{sample}}}{(E\ \text{target})^{Ct\ \text{sample}}} \div \frac{(E\ \text{ref})^{Ct\ \text{control}}}{(E\ \text{target})^{Ct\ \text{control}}}$$

Where E= qRT PCR efficiency, Ref= internal control gene, Target= target gene, Ct = Cycle threshold value and Δ = Change in Ct value of target gene with respect to internal control gene

Applications of qRT-PCR:

qRT-PCR can be applied to traditional PCR applications as well as new applications that would have been less effective with traditional PCR. With the ability to collect data in the exponential growth phase, the power of PCR has been expanded into applications such as: Viral Quantitation, Quantitation of Gene Expression, Array Verification, Drug Therapy Efficacy, DNA Damage measurement, Quality Control and Assay Validation, Pathogen detection and Genotyping.

FLUORESCENCE IN SITU HYBRIDIZATION (FISH)

FISH is used to detect small deletions and duplications that are not visible for microscope analysis. It is used to detect the presence of no. of chromosomes of a certain type in each cell or to substantiate rearrangements that are visible in microscope analysis. FISH is used specifically to look at a specific area of one chromosome only. It uses A very small chemical is used which glows brightly after detecting a specific region on the chromosome. A special microscope is used to look at the chromosome to find out the no. of bright spots. In the case of deletion, only one bright spot is visible instead of two (one on each chromosome) whereas in the case of a duplication, three bright spots are found instead of just two.

FISH uses probes, small DNA strands, which have a fluorescent label attached to them. These probes are designed to be complementary to specific parts of a chromosome. When DNA is denatured on heating, the probes hybridise (Fig 7) to their complementary sequence in the DNA of the patient. The probe will not hybridise if a small deletion is found in the region complementary to the probe. More no. of probe hybridise if duplication is present.

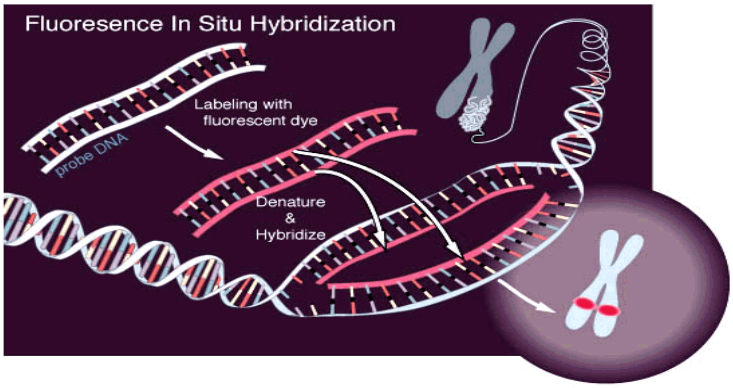


Fig.7: Hybridization of probes in FISH Procedure

These are the following steps in the FISH procedure (Fig. 8):

i) Denaturing of the chromosomes
ii) Denaturing of the probe
iii) Hybridization
iv) Fluorescence staining
v) Examining the slides or storing in the dark

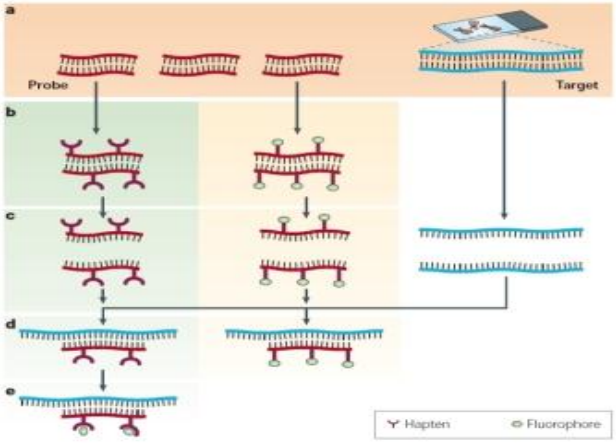


Figure 8: Schematic diagram for FISH technique.Reprinted from O'Connor, 2008. Fluorescence *in situ* hybridization (FISH)

FISH testing for a deletion

Two types of probes are generally used; the first probe (generally green) is known as control probe which identifies both copies of the chromosome under test. It gets hybridised to a sequence that is not in the deletion region, so a signal is seen on each chromosome. The second probe (generally red) gets hybridised to the sequence that may be deleted. A deletion is generally present in only one of the chromosomes in a pair, therefore the probe bind to the undamaged

chromosome, but is incapable to bind to the deleted chromosome which results in only one signal.

Types of probes

Three different types of FISH probes are used. Each of them has a different application:

Locus specific probes – These probes bind to a particular region of a chromosome and is used when a small portion of a gene is isolated to determine on which chromosome the gene is located.

i) **Centromeric repeat probes** – Alphoid or centromeric repeat probes are the result of repetitive sequences present in the middle of each chromosome. These probes are used to determine whether an individual has the correct number of chromosomes. These probes are also used in combination with locus specific probes for determining missing genetic material from a particular chromosome in an individual.

ii) **Whole chromosome probes** – These are actually collections of smaller probes. Each of the smaller probes binds to a different sequence along the length of a given chromosome. They are particularly useful for examining chromosomal abnormalities viz. when a piece of one chromosome is found to be attached to the tail of another chromosome.

- **Samples for FISH testing:**

  FISH is mostly performed on blood samples from both, adults and children. FISH is also used as a prenatal test for aneuploidy, extra copies of whole chromosomes, using placental samples from chorionic villus sampling (CVS) or amniotic fluid from an amniocentesis. It is also used as a prenatal test for deletions.

- **Limitations of FISH:**
  FISH probes are available for the well characterized duplication and deletion syndromes only. It is difficult to see a duplication using FISH because the attachment of extra probes is not very easy to observe.

Applications of FISH in Clinical Practice

i) Pre-implantation Diagnostics
ii) Prenatal Diagnostics
iii) Tumor Diagnostics
iv) Postnatal Diagnostics
v) Centromere Mutations

DNA SEQUENCING TECHNIQUES

Sequencing means to find the order of nucleotides on a string of DNA. A variation in a DNA sequence can lead to a changed or non-functional protein, and lead to a genetic disorder. Sequencing of DNA is important to identify the type of mutations in genetic diseases.

The four best known DNA sequencing techniques are:

i) Sanger method and its most important variants (enzymic methods);
ii) Maxam & Gilbert method and other chemical methods;
iii) Pyrosequencing™ method – DNA sequencing in real time by the detection of released pyrophosphate (PPi) ; and
iv) Single molecule sequencing with exonuclease (exonuclease digestion of a single molecule composed of a single strand of fluorescently labelled deoxynucleotides).

i) Sanger's method and other enzymatic methods

Initially known as the chain termination method or the dideoxynucleotide method, it is comprised of a catalysed enzymatic reaction that polymerizes the DNA fragments complementary to the

template DNA of interest i.e. unknown DNA. A ^{32}P-labelled primer, short oligonucleotide with a sequence complementary to the template DNA, is annealed to a precise known region on the template DNA, which provides an initial point for DNA synthesis. Catalytic polymerization of deoxynucleoside triphosphates (dNTP) occurs onto the DNA in the presence of DNA polymerases. The polymerization is extended and the enzyme incorporates a modified nucleoside into the growing chain. Different approaches, shotgun (random) and primer walking (direct) sequencing mostly, are used and their strategies are described below in detail.

Random approach

In the usual procedure, with few exceptions, there is no control of the region that is going to be sequenced. It is also known as shotgun sequencing. It has ready availability of optimized cloning vectors, fluorescently labelled universal primers also and software for the purpose of base calling and sequence assembly. The whole process comprises high level of automation, from the cloning of the vectors and colony selection until the bases called.

Direct approach

In direct sequencing, the sequence is known of unknown DNA. This approach is also known as primer walking (Fig. 9). Its major advantage is the reduced redundancy because of the direct nature of the approach which is opposite of random approach. However, this approach requires the synthesis of each new primer, which is expensive.

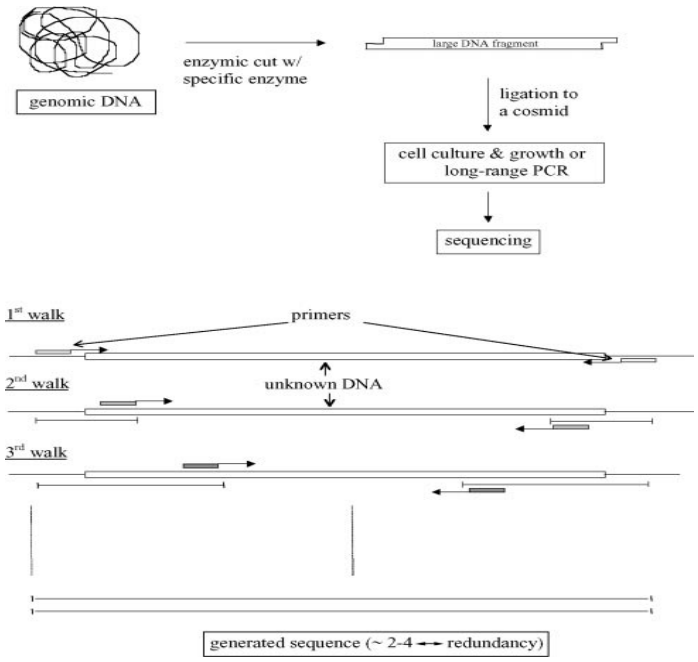


Figure 9: DNA sequencing by the primer-walking strategy

The genomic DNA is cut into a large piece (~ 40 kbp) and inserted into a cosmid for growth. The sequencing is performed by walks i.e. starting first from the known region of the cosmid. After the results from the first round are edited, a new priming site is located within the newly generated sequence. This procedure is repeated and gets over only when walks reach the opposite starting points.

Enzyme technology

Developments in DNA polymerase enzymes have significantly helped in the quality of the sequencing reactions and sequencing data. In fluorescently labelled dye-terminators, there was significant variation in peak intensity. The system of the termination was reproducible and predictable, but this deviation made automatic base calling difficult. A

few years later, a modified set of fluorescent labels for ddNTPs are introduced. The signal was more even, and automated base calling improved significantly.

Sample preparation

The following steps are often included in the methodology for sample preparation: (i) DNA scission and cloning into a vector (e.g. M13 or M13mp18); (ii) vector amplification to produce a phage-infected culture; and (iii) purification from the cell culture to yield pure single-stranded (ss) DNA template.

Labels and DNA labelling

A) Radioisotopes

The enzymic method used ^{32}P as a label in the iitial stage. Biggin *et al*. (1983) suggested the use of deoxyadenosine 5'-(α-[^{35}S] thio) triphosphate as the label unifiedinto the DNA fragments. This strategy stemmed in intensification of the band sharpness on autoradiography and the resolution of band separation.

B) Chemiluminescent detection

Chemiluminescent detection is used as an alternative to radioisotopes, a method based on Biotin – streptavidin system. The 5'-end of an oligonucleotide linked to biotin is used as the primer in the sequencing reaction. There are three major advantages to this method; first, the sequencing reactions are obtained directly from the PCR products; secondly, this method does not require cloning of the DNA before sequencing, and thirdly, it is feasible to multiplex several reactions on the same gel and detect one at a time with suitable enzyme linked primers.

C) Fluorescent dyes

Smith *et al*. (1986) developed a set of four different fluorescent dyes that allowed four reactions to be separated in a single lane.

Alternative dyes are synthesized. These dyes have emission spectra with their maxima relatively well spaced, which simplifies colour/base discrimination. One drawback of this group of dyes is the need for two wavelengths for excitation; for FAM and JOE dyes at 488 nm, and for TAMRA and ROX dyes at 543 nm.

ii) Maxam & Gilbert and other chemical methods

In this method, described by Maxam & Gilbert (1977), end-labelled DNA fragments are exposed to random cleavage at adenine, cytosine, guanine, or thymine positions by using specific chemical agents. The chemical attack is divided on three steps: base modification, removal of the modified base from its sugar, and DNA strand breaking at that sugar position. The products of these four reactions are then separated using polyacrylamide gel electrophoresis.

The chemical reactions can be separated into two different groups: (i) four-lane methods, where four (or more) separate cleavage procedures are used and the information is displayed in four (or more) parallel gel lanes and (ii) one-lane (or two-lane) method, where all reactions are based on only one chemical modification and electrophoresis is performed in a single (or two) lane(s).

iii) Pyrosequencing – DNA sequencing in real time by the detection of released PPi

Pyrosequencing, a real-time DNA-sequencing method, is based on the detection of the PPi released during the DNA polymerization reaction. Initially, this approach was used for continuous monitoring of DNApolymerase activity. Presently, there are two different pyro sequencing approaches: solid-phase sequencing and liquid-phase sequencing. The main problem noticed in all versions of pyrosequencing techniques is the interference of dATP in the detection of luminescence. This problem is solved by replacing dATP by dATPαS in the experimental step.

Pyrosequencing has shown several advantages:

i) Detection is in real time with a cycle time of approximately 2 min (solid-phase) ;
ii) Reactions occur at room temperature and physiological pH;
iii) Method is easily adapted for multiplexed sample processing ;

On the other hand, this method has presented some disadvantages such as:

i) In the solid phase approach, the template must be washed completely after each nucleotide addition, resulting in a decreased signal due to loss of templates ;
ii) In the liquid phase approach, the apyrase activity is decreased in later cycles due to accumulation of intermediate products.
iii) Contamination with PPi decreases the signal-to-noise ratio significantly, due to increased background signal.

The main applications of this method are *de novo* DNA sequencing for short- and medium-length DNA, analysis of single-nucleotide polymorphisms, mutation detection and the analyses of secondary structure, such as hairpin structures.

Single-molecule sequencing with exonuclease

It was initially conceived as a laser-based technique allowing the fast sequencing of DNA fragments of 40 kb or more at a rate of 100 – 1000 bases/second. This technique is based on the identification of individual fluorescent nucleotides in a flowing sample stream. The method is divided into the following steps (Fig. 10): fluorescent labelling of the bases in a single fragment of DNA, attachment of this labelled DNA fragment onto a microsphere, movement of the supported DNA fragment into a flowing buffer stream, digestion of the labelled DNA with an exonuclease that sequentially cleaves the 3'-end nucleotides, and

detection and identification of individual fluorescently labelled bases as they cross a focused laser beam.

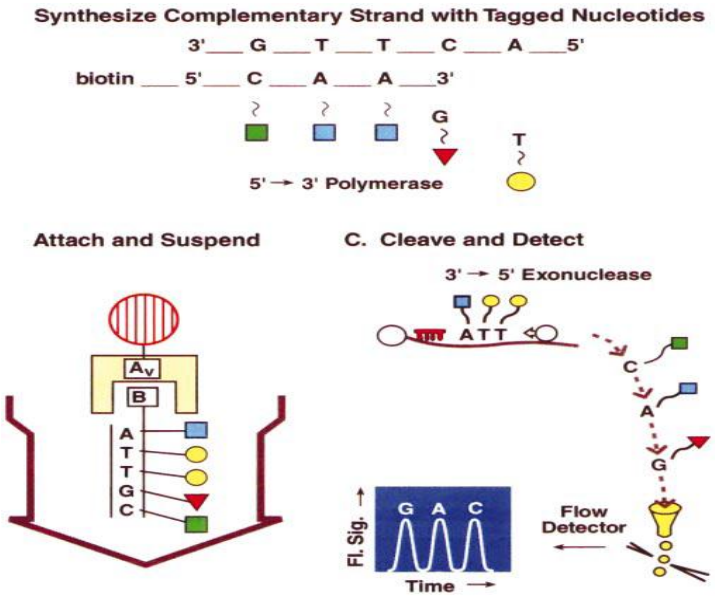


Figure 10: Schematic representation of the steps involved in a single molecule sequencing experiment with a detection setup based on laser induced fluorescence. (Reproduced from Davis et al. 1991)

There are still many problems that have to be solved. The quality of buffer must be improved. A selection step must be incorporated into the sequencing process. Complementary DNA strands have to be labelled with four different nucleotides. New polymerases and new exonucleases are necessary for rapid and efficient sequencing.

MICROARRAY

Microarray technology is used to monitor genome wide expression levels of genes in a known organism. In this technology, a

glass slide is used on to which DNA molecules are fixed in an organized manner at definite locations called spot. It may contain thousands of spots and these spots may hold a few million copies of identical DNA molecules that exclusively belong to a gene (Fig. 11A). The DNA in a spot may either be genomic DNA or short stretch of oligonucleotide strands that correspond to a gene. The spots or features are synthesised by the process of photolithography or printed on to the glass slide by a robot.

Microarrays are used to measure gene expression in several ways. One of the several popular applications is to compare the expression of a set of genes from a cell, to be diagnosed, (condition A) to the same set of genes from a reference cell maintained under standard conditions (condition B). Figure 11B provides a general picture of the involved experimental steps. First, RNA is extracted from the cells and reverse transcribed into cDNA by using reverse transcriptase enzyme and nucleotides labelled with different fluorescent dyes. cDNA from cells developed in condition A may be labelled with a red dye and from cells developed in condition B with a green dye. After differentially labelling the samples, they are hybridized onto the same glass slide. Any cDNA sequence in the sample hybridizes to specific spots present on the glass slide with its complementary sequence. The amount of cDNA bound to a spot is found to be directly proportional to the initial no. of RNA molecules available for that gene in both the samples.

The spots in the hybridized microarray are made to excite by a laser and scanned at appropriate wavelengths to distinguish the red and green dyes. The amount of fluorescence emitted upon excitation corresponds to the amount of nucleic acid bound. If cDNA from condition A for a particular gene is in greater quantity than that from condition B, the spot is found to be red. If cDNA from condition B is in abundance, the spot would be green. If the gene is expressed equally in both conditions, yellow spot is found, and if the gene is not expressed in

either condition, black spot is found. Thus, the end of the experiment is an image of the microarray where each spot corresponding to a gene has an associated fluorescence value which represents the relative expression level for that gene.

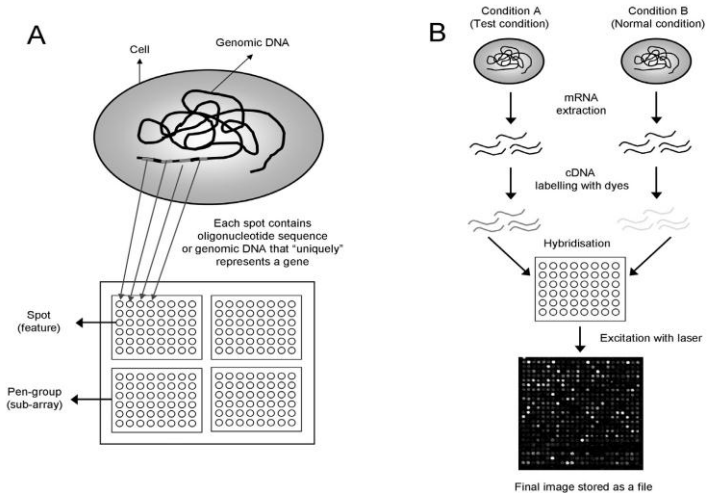


Figure 11(A) A microarray may contain thousands of spots. Each spot contains many copies of the same DNA sequence that uniquely represents a gene from an organism. Spots are arranged in an orderly fashion. (B) Schematic of the experimental protocol to study differential expression of genes

The organism is grown in two different conditions (a reference condition and a test condition). RNA is extracted from the two cells, and is labelled with different dyes (red and green) during the synthesis of cDNA by reverse transcriptase. This step is followed by hybridization of cDNA onto the microarray slide, where each cDNA molecule representing a gene will bind to the spot which contains its complementary DNA sequence. The microarray slide is then exposed to a laser for excitement at suitable wavelengths to detect the red and green dyes. The final image obtained is stored as a file for additional study.

Application of Microarrays

- Sensitive enough to detect single base differences, SNPs or mutations
- Useful for a wide range of applications viz. identifying contamination of food products with cells from other animals or plants; identifying strains of viruses; detecting mutations in cancer cells of a patient that may influence the disease's response to treatment.
- Protein microarrays are developed for massive screening of interactions between proteins on the microarray, and other proteins, ligands or substrates.

NEXT GENERATION SEQUENCING (NGS)

Nucleic acid sequencing helps in determining the order of nucleotides found in a DNA or RNA molecule. The use of nucleic acid sequencing has increased exponentially. The first main venture into DNA sequencing was the Human Genome Project, a $3 billion endeavour which took 13-year-long period to be completed in 2003. The Human Genome Project was performed with Sanger sequencing, first-generation sequencing. Sanger sequencing, the chain-termination method, was developed in 1975 by Edward Sanger and was used for nucleic acid sequencing for the successive two and a half decades (Sanger et al., 1977).

Demand for cheaper and faster sequencing methods has enhanced greatly since the completion of the first human genome sequence which lead to the development of second-generation sequencing methods, or next generation sequencing (NGS). NGS platforms execute enormous parallel sequencing and millions of fragments of DNA from a single sample are sequenced in unison. Massive parallel sequencing technology facilitates an entire genome to

be sequenced in less than one day. In the previous decade, numerous NGS platforms have been developed to provide low-cost and high-throughput sequencing. The Illumina MiSeq, The Life Technologies Ion Torrent Personal Genome Machine (PGM) and other NGS platforms has made sequencing available to more no. of labs. Resultantly, the volume of research and clinical diagnostics being performed with the help of nucleic acid sequencing is rapidly increasing.

Overview of the methodology

The Ion Torrent PGM and the Illumina MiSeq have a common base methodology that includes template preparation, sequencing and imaging, and data analysis. An overview of thesequencing methodologies is provided in Figure 12.

Template preparation

This step consists of construction of a library of DNA or complementary DNA and amplification of that library. Sequencing libraries are constructed by fragmenting the DNA or cDNA sample and ligating adapter sequences i.e. synthetic oligonucleotides of a known sequence onto the ends of DNA fragments. After construction, libraries are clonally amplified in preparation for the purpose of sequencing. The MiSeq utilizes bridge amplification for the formation of template clusters on a flow cell whereas the PGM utilizes emulsion PCR on the OneTouch system to increases single library fragments onto microbeads,

Sequencing and imaging

To acquire nucleic acid sequence from the amplified libraries, the MiSeq and the Ion Torrent PGM both rely on sequencing by the process of synthesis. The library fragments are used as template on which a new DNA fragment is synthesized. The sequencing comprises a cycle of washing and flooding of the fragments with the known nucleotides in a sequential order. As nucleotides integrate into the

increasing DNA strand, they are digitally recorded as sequence. The MiSeq and the PGM rely on a somewhat altered mechanism for detecting nucleotide sequence information. The MiSeq depends on the detection of fluorescence produced by the incorporation of fluorescently labelled nucleotides into the growing strand of DNA. By contrast, the PGM performs semiconductor sequencing that depends on the detection of pH changes caused by the release of a hydrogen ion upon the incorporation of a nucleotide into a growing strand of DNA.

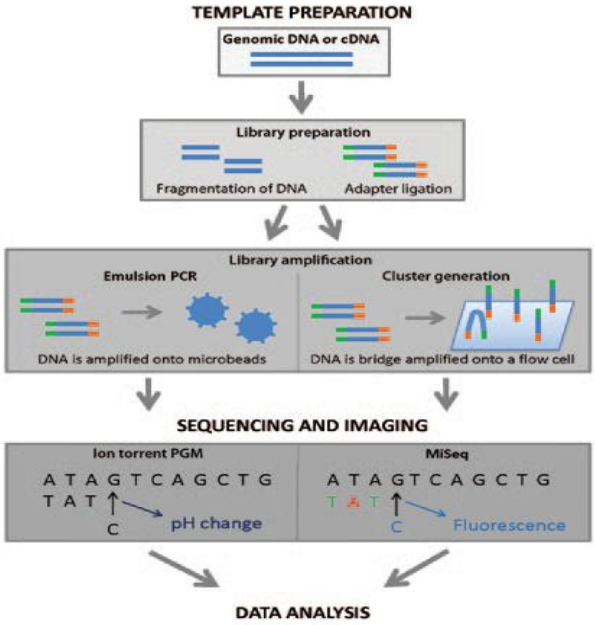


Figure 12: Next-generation sequencing methodology

Data analysis

After completion of sequencing, raw sequence data undergo various analysis steps. A general data analysis for NGS data includes pre-processing of the data for removing adapter sequences and low-quality

reads, the mapping of the data to a reference genome. Analysis of the sequence includes a wide variety of assessments using bioinformatics, including detection of novel genes or regulatory elements, genetic variant calling for detection of SNPs or indels i.e. the insertion or deletion of bases and assessment of transcript expression levels. Analysis also includes the identification of somatic and germline mutation events that contributes to the correct diagnosis of a disease or genetic condition.

Applications

- Helps in comparative biology studies through whole genome sequencing of a wide variety of organisms.
- Sequencing of the human genome is being performed again to identify genes and regulatory elements implicated in pathological processes.
- Gene expression studies using RNA-Seq (NGS of RNA) have started replacing the use of microarray analysis which provides researchers the ability to analyse RNA expression in sequence form.
- In the fields of public health and epidemiology through the sequencing of bacterial and viral species to facilitate the identification of novel virulence factors.

NGS in practice

A) Whole-exome sequencing

Exome sequencing is used extensively in the past several years for gene discovery research. Some gene discovery studies have resulted in the documentation of genes that are relevant to inherited skin diseases. It can also assist in the identification of disease-causing mutations where the particular genetic cause is not known.Figure 13

establishes the direct effect that NGS in the correct diagnoses of a patient. It explains the practise of homozygosity mapping followed by whole-exome sequencing to recognize two disease- causing mutations in a patient with oculocutaneous albinism and congenital neutropenia. Figure 13a and 13b show the phenotypic traits common to oculocutaneous albinism type 4 and neutropenia detected in this patient. Figure 13c is a pedigree of the family of a patient, both the affected and unaffected individuals. The ideogram i.e. the graphic chromosome map in figure 13d stresses the regions of genetic homozygosity. These areas were acknowledged by single-nucleotide-polymorphism array analysis and were believed to be probable locations for the disease-causing mutation(s). Figures 13e and 13f exhibit chromatograms for the two disease-causing mutations recognised by whole-exome sequencing. Figure 13e portrays the mutation in *SLC45A2,* and Figure 12f shows the mutation in *G6PC3*. This case depicts the valuable role that NGS can play in the precise analysis of an individual patient who shows different symptoms with an unrevealed genetic cause.

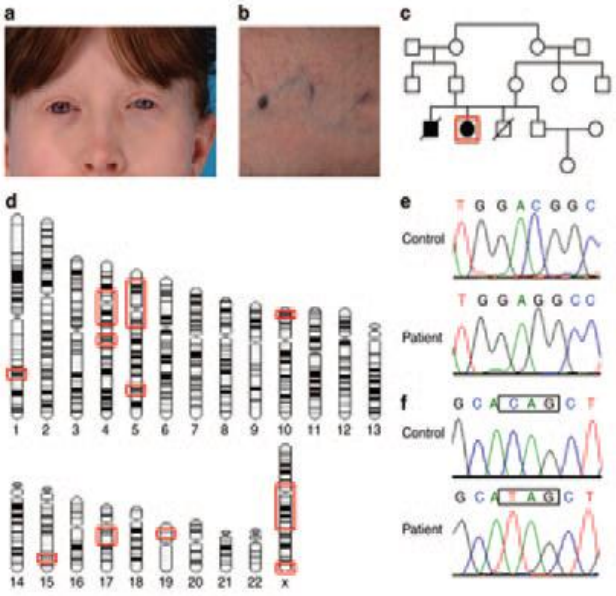


Figure 13.Clinical application of whole-exome sequencing in the detection of two disease-causing mutations. Reprinted from Cullinane *et al*., 2011

B) Targeted sequencing

Targeted sequencing of specific genes or genomic regions is preferred when a suspected disease or condition has been identified. It is more affordable, reduces sequencing cost and time, and yields much higher coverage of genomic regions of interest. Sequencing panels are being developed for targeting hundreds of genomic regions known for hotspots for disease-causing mutations. Since, these sequencing panels target desired regions of the genome for sequencing; it eliminates the majority of the genome from analysis. The results of targeted sequencing of diseases help in decision making in many diseases therapeutically which includes different cancers for which the treatments should be cancer-type specific preferably.

Limitations

NGS is still too expensive for many lab conditions. Erroneous sequencing of homopolymer regions (spans of repeating nucleotides) on certain NGS platforms and short-sequencing read lengths (on average 200–500 nucleotides) can lead to sequence errors. Data analysis is sometimes time-consuming and requires knowledge of bioinformatics to harvest precise information from sequence data.

References:

1. Ana, L. R., Miguel, Gómez-Lim b, Francisco Fernández a, Achim M. Loske., Physical methods for genetic plant transformation (2012) Physics of Life Reviews 9: 308–345.
2. Ayman Grada and Kate Weinbrecht (2013). Next-Generation Sequencing: Methodology and Application.Jour. of Investigative Dermat. 133,1-4.
3. Biggin, M.D., Gibson, T. J. & Hong, G. F. (1983). Buffer gradient gels and $&S label as an aid to rapid DNA sequence determination. Proc. natn. Acad. Sci. USA 80, 3963-3965.
4. Birnboim, H. C. and Doly, J. A., 1979, Rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7 (6): 1513-1523.
5. **Bustin**, S. A., (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. J. mol. Endocrinol. 29(1): 23-39.
6. Causton, H., Quackenbush, J., and Brazma, A. (2003). Microarray Gene Expression Data Analysis: A Beginner□s Guide. (UK: Blackwell Science).
7. Cho, R. J., Campbell, M. J., Winzeler, E. A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T. G., Gabrielian, A. E., Landsman, D., Lockhart, D. J., and Davis, R. W. (1998). A genome-wide transcriptional analysis of the mitotic cell cycle. Mol. Cell. 2, 65-73.
8. Czechowski, T., Bari, R. P., Stitt, M., Scheible, W. R. and Udvardi, M. K., (2004) Real-time RT-PCR profiling of over 1400 *Arabidopsis* transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. Plant J., 38: 366-379.
9. Davis, L. M., Fairfield, F. R., Harger, C. A., Jett, J. H., Keller, R. A., Hahn, J. H., Krakowski,

a. L. A., Marrone, B. L., Martin, J. C., Nutter, H. L., Ratliff, R. L., Shera, E. B., Simpson, D. J. & Soper, S. A. (1991). Rapid DNA sequencing based upon single molecule detection. Genetic Analysis – Biomolec. Engng 8, 1–7.

10. Diatchenko, L., Yun-fai, C.L., Aaron, P., Campbell, A.C., Fauziamoqadam, B.H., Sergey, L., Konstantin, L., Nadya, G., Eugene, D.S. and Paul, D.S., (1996) Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci.*, USA, 93: 6025-6030.
11. Dopazo, J., Zanders, E., Dragoni, I., Amphlett, G., and Falciani, F. (2001). Methods and approaches in the analysis of gene expression data. J. Immunol. Meth. 250, 93-112.
12. Doyle, J. J. and Doyle, J. L, (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull., 19: 11–15.
13. Hegedas, D. D. and Khazhatourians, G. G., (1996) Detection of entomopathogenic fungus *Bauveria bassiana* within infected migratory grasshoppers (*Melanoplus sanguipiper*) using polymerase chain reaction and DNA probe. J. Invertebrate Pathol., 67: 21-27.
14. Kumar, P., Gupta, V. K., Misra, A. K., Modi, D. R. and Pandey, B. K., (2009) Potential of Molecular Markers in Plant Biotechnology. Plant Omics Journal.2(4):141-162.
15. Lilian, T. C. Franc:a, Emanuel Carrilho and Tarso, B. L. Kist. (2002). A review of DNA sequencing techniques. Quarterly Reviews of Biophysics 35 (2),169–200.
16. Marilena Aquino de Muro., (2005) Probe Design Production and Applications, Molecular Biomethods. John M. Walker, Ralph Rapley (edited). Second edition. Humana press. 41-54.
17. Maxam, A.M. & Gilbert, W. (1977). A new method for sequencing DNA. Proc. natn. Acad. Sci. USA 74, 560-564.

18. Michels, C. A., (2002) Gene isolation and analysis of multiple mutant alleles. Genetic Techniques for Biol. Res. pp 85-90. John Wiley & Sons, Ltd publication.
19. Pfaffl, M.W., (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res., 29: 45-51.
20. Ramakers, C., Ruitjer, J. M., Deprez, R. H. and Moorman, A. F., (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci. Lett., 13: 62-66.
21. Sambrook, J. and Russell, D. W., (2001) Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring harbor Laboratory Press, cold Spring Harbor, NY.
22. Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–7
23. Semagn, K., Bjørnstad, Å. and Ndjiondjop, M. N., (2006) Principles, requirements and prospects of genetic mapping in plants. African Journal of Biotechnology Vol. 5 (25), pp. 2569-2587.
24. Singh, B. D., (2010) Biotechnology Expanding Horizons . Kalyani Publishers. pp. 33-34.
25. Smith, L.M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S. B. H. & Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. Nature 321, 674-679.
26. Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology 98: 503–517.
27. Terence, A. B., (2001) Southern Blotting and Related DNA Detection Techniques. Encyclopedia of Life Sciences. 2001,
28. Tichopad, A., Dilger, M., Schwarz, G. and Pfaffl, M. W., (2003) Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic Acids Res*., 31: 122.
29. Wilcox, J. (1993) Fundamental principles of in situ hybridization. J. Histochem. Cytochem. 41, 1725-1733.

Chapter 9
Antibiotics: Microbial Sources, Production and Optimization

Ranjith Kumar.M, Ashokraj.S and Brindha Priyadarisini.V

1. Antibiotics

Antibiotics are a group of natural or synthetic compounds that destroy bacteria (bactericidal) or inhibit their growth (bacteriostatic). Antibiotics that are sufficiently nontoxic to the host are used as chemotherapeutic agents in the treatment of infectious diseases of humans, and animals. Nature produces an amazing variety and number of products. In this section we will concentrate on antibiotics and its natural sources. About 100,000 secondary metabolites of molecular weight less than 2500 have been characterized, which are mainly produced by microbes and plants (Roessner and Scott, 1996); Out of which around 50,000 are from microorganisms (Fenical and Jensen, 1993; Berdy, 1995).

The selective action exerted on pathogenic bacteria and fungi by microbial secondary metabolites ushered in the antibiotic era, and for 50 years we have been benefited from this remarkable property of “wonder drugs” such as penicillins, cephalosporins, tetracyclines, aminoglyco sides, chloramphenicol, and macrolides, among others. The successes were so impressive that these antibiotics were virtually the only drugs utilized for chemotherapy against pathogenic microorganisms. By 1996,

the world market for antimicrobials amounted to $23 billion and involved some 150 to 300 products, natural, semisynthetic, or synthetic. The $8 billion US antimicrobial market in 1995 included cephalosporins (45%), penicillins (15%), quinolones (11%), tetracyclines (6%) and macrolides (5%) (Strohl, 1997).

About 20 years ago, the difficulty and high cost of isolating novel structures and antimicrobial agents with new mode of action for such uses became apparent, and the field looked like it might enter a phase of decline. Indeed, the number of anti-infective investigational new drugs (INDs) declined by 50% from the 1960s to the late 1980s (DiMasi*et al*., 1994). However, it was realized that compounds which possess antibiotic activity also possess other activities, that some of these had been quietly exploited in the past, and that such broadening of scope should be exploited and expanded in the future. Thus, a broad screening of antibiotically active molecules for antagonistic activity against pathogenic organisms other than microorganisms, as well as for other pharmacological applications, was proposed in order to yield new and useful lives for "failed antibiotics". A large number of *in vitro* laboratory tests were developed to help detect, isolate, and purify useful compounds. Much of this emphasis was brought about by Umezawa (1982), who pointed out the potential importance of enzyme inhibitors as drugs. Fortunately, we entered into a new era in which microbial metabolites were applied to diseases i.e., diseases not caused by bacteria and fungi. Let us see some of the bacterial antibiotics in detail in this chapter.

1.2. Antibacterial agents

The fight against bacterial infection is one of the great success stories of medicinal chemistry. During that latter half of the nineteenth century, scientists such as Koch were able to identify the microorganisms

responsible for diseases such as tuberculosis, cholera, and typhoid. Methods such as vaccination for fighting infections were studied. Research was also carried out to try and find effective antibacterial agents or antibiotics. However, the scientist who can lay claim to be the father of chemotherapy the use of chemicals against infection was Paul Ehrlich. Ehrlich spent much of his career studying histology, then immunochemistry, and won a Nobel prize for his contributions to immunology. However, in 1904 he switched direction and entered a field which he defined as chemotherapy. Ehrlich's 'Principle of Chemotherapy' was that a chemical could directly interfere with the proliferation of microorganisms, at concentrations tolerated by the host. This concept was popularly known as the 'magic bullet', where the chemical was seen as a bullet which could search out and destroy the invading microorganism without adversely affecting the host. The process is one of selective toxicity, where the chemical shows greater toxicity to the target microorganism than to the host cells. Such selectivity can be represented by a 'chemotherapeutic index', which compares the minimum effective dose of a drug with the maximum dose which can be tolerated by the host. This measure of selectivity was eventually replaced by the currently used therapeutic index. By 1910, Ehrlich had successfully developed the first example of a purely synthetic antimicrobial drug. This was the arsenic-containing compound salvarsan. Although it was not effective against a wide range of bacterial infections, it proved effective against the protozoal disease sleeping sickness (trypanosomiasis), and the spirochaete disease of syphilis. The drug was used until 1945 when it was replaced by penicillin.

Over the next twenty years, progress was made against a variety of protozoal diseases, but little progress was made in finding antibacterial agents, until the introduction of proflavine in 1934. Proflavine is a yellow colored amino acridine structure which is

particularly effective against bacterial infections in deep surface wounds, and was used widely during the Second World War. It is an interesting drug since it targets bacterial DNA rather than protein. Despite the success of this drug, it was not effective against bacterial infections in the bloodstream and there was still an urgent need for agents which would fight these infections. This need was answered in 1935 with the discovery that a red dye called prontosil which was effective against *Streptococci* infections *in vivo.* Prontosil was eventually recognized as being a prodrug for a new class of antibacterial agents such as the sulfa drugs (sulfonamides). The discovery of these drugs was a real breakthrough, since they represented the first drugs to be effective against bacterial infections carried in the bloodstream. They were the only effective drugs until penicillin became available in the early 1940s.

Although penicillin was discovered in 1928, it was not until 1940 that effective means of isolating it were developed by Florey and Chain. Society was then rewarded with a drug which revolutionized the fight against bacterial infection and proved even more effective than the sulfonamides. Despite penicillin's success, it was not effective against all types of infection and the need for new antibacterial agents still remained. Penicillin is an example of a toxic chemical produced by a fungus to kill bacteria which might otherwise compete with it for nutrients. The realization that fungi might be a source for novel antibiotics spurred scientists into a huge investigation of microbial cultures, both known and unknown.

In 1944, the antibiotic streptomycin was discovered from a systematic search of soil organisms. It extended the range of chemotherapy to *Tubercle bacillus* and a variety of Gram-negative bacteria. This compound was the first example of a series of antibiotics known as the aminoglycoside antibiotics. After the Second World War, the effort continued to find other novel antibiotic structures. This led to

the discovery of the peptide antibiotics (e.g. bacitracin (1945)), chloramphenicol (1947), the tetracycline antibiotics (e.g Chlortetracyc line (1948)), the macrolide antibiotics (e.g. erythromycin (1952)) and the cyclic peptide antibiotics (e.g. cycloserine (1955)). As far as synthetic agents were concerned, isoniazid (a pyridine hydrazide structure) was found to be effective against human tuberculosis in 1952, and in 1962 nalidixic acid the first quinolone antibacterial agents was discovered. A second generation of this class of drugs was introduced in 1987 with ciprofloxacin. Many antibacterial agents are now available and the vast majority of bacterial diseases have been brought under control (e.g. syphilis, tuberculosis, typhoid, bubonic plague, leprosy, diphtheria, gas gangrene, tetanus, gonorrhea). This represents a great achievement for medicinal chemistry and it is perhaps sobering to consider the hazards which society faced in the days before penicillin.

Of the 12,000 antibiotics known in 1995, 55% were produced by filamentous bacteria (actinomycetes) of the genus Streptomyces, 11% from other actinomycetes, 12% from non filamentous bacteria and 22% from filamentous fungi (Berdy, 1995; Strohl, 1997).

2. Actinomycetes

The phylum actinomycetes are one of the largest groups in the domain bacteria largely consists of environmental bacteria and the denizens of many varied habitat soils such as the rhizosphere, marine and extreme arid environments. Actinomycetes typically have elevated guanosine-cytosine contents (65-75% G + C) and their genome sizes range from 2.5-Mb skin commensal *Micrococcus luteus* to the 9.7-Mb environmental strain *Rhodococcusjostii*. Since the discovery of antibiotics in the 1940s, the actinomycetes have received a great deal of attention, and *Streptomyces* species in particular have become renowned as the principal sources of therapeutic pharmaceuticals. There have been

several good reviews on actinomycetes of late, notably that by Ventura *et al*. (2007) on evolutionary and genomic aspects, as well as occasional articles focusing on specific genera.

However, other genera, including *Rhodococcus*, are beginning to excite more interest and Streptomyces may command less attention in the future (Larkin *et al*., 2005, Kitagawa and Tamura, 2008). Streptomycetes are demonstrably a rich source of compounds, but no more so than other members of the actinomycetes, also the Bacilli and bacteria genera such as the myxobacteria (Wenzel and Muller, 2009) and pseudomonas (Gross and Loper, 2009). Among the eukaryotes, fungal genomes are replete with biosynthetic gene clusters for encoding small molecule production. The ability to make bioactive small molecules is not exclusive to microbes. There is a global crisis in the treatment of infectious diseases and people are dying of infections that were previously treatable. Microbes are the source and the solution for the crisis, and for this reason it is imperative that the search for novel therapeutic agents be intensified. The constant moan of the pharmaceutical industry, that the natural reservoir of molecules with antibiotic activity is close to being exhausted and that they can find no useful bioactive compounds. It can also be explained by the inability to detect bioactive compounds when they are present only in low concentrations. The industry has found all the easily accessible bioactive called ‘‘low hanging fruit’’ (Baltz, 2006). Presumably it was not considered essential to develop the technology to find compounds that were missed.

In addition, actinomycetes as a whole have been ignored in the past years even though they too possess the capacity to produce a huge number of bioactive small molecules. Till date, only a small proportion of actinomycetes have been examined for therapeutic purposes. We are now in the ‘‘genomic era’’ and in the case of Streptomycetes, exciting

new information coming from complete genome sequencing efforts reveals that most of these bacteria have the genetic capacity to produce many more structurally different bioactive compounds than suspected. As such, they represent an inexhaustible collection of hidden chemical and biochemical diversity. Moreover, creative techniques for generating some of these compounds are being developed and exploited (Baltz, 2008; Challis, 2008).

2.1. Marine actinomycetes

Actinomycetes represent a ubiquitous group of microbes widely distributed in natural ecosystems around the world and especially significant for their role in the recycling of organic matter (Srinivasan *et al*., 1991). Actinomycetes are abundant in terrestrial soils, a source of most isolates shown to produce bioactive compounds. Goodfellow and Haynes (1984) reviewed the literature on the isolation of actinomycetes from marine sediments and suggested that this source may be valuable for the isolation of novel actinomycetes with the potential to yield useful new products. Earlier studies (Weyland, 1981; Helmke and Weyland, 1984) considered actinomycetes to be part of an indigenous marine microflora. Others saw them primarily as wash in components that nearly survived in marine and littoral sediments as spores (Goodfellow and Williams, 1983). Jensen *et al*. (1991) and Takizawa *et al*. (1993) reported a bimodal distribution of actinomycetes in near shore tropical marine environments. The existence of an autochthonous actinomycete population was suggested by the isolation of actinomycetes from marine deep oceanic sediments (Takizawa *et al*., 1993; Ravel *et al*., 1998; Weyland and Helmke, 1988).

Although the diversity of life in the terrestrial environment is extraordinary, the greatest biodiversity is in the oceans. More than 70% of our planet's surface is covered by oceans and life on earth originated from the sea. In some marine ecosystems, such as the deep sea floor and

coral reefs experts estimate that the biological diversity is higher than in the tropical rainforests. As marine environmental conditions are extremely different from terrestrial ones, it is surmised that marine actinomycetes have different characteristics from those of terrestrial counterparts and, therefore, might produce different types of bioactive compounds. The living conditions to which marine actinomycetes had to adapt during evolution range from extremely high pressure and anaerobic conditions at temperatures just below 0˚C on the deep sea floor, to high acidic conditions (pH as low as 2.8) at temperatures of over 10˚C near hydrothermal vents at the mid-ocean ridges. It is likely that this is reflected in the genetic and metabolic diversity of marine actinomycetes, which remains largely unknown. Indeed, the marine environment is a virtually untapped source of novel actinomycetes diversity and, therefore, of new metabolites. However, the distribution of actinomycetes in the sea is largely unexplored and the presence of indigenous marine actinomycetes in the oceans remains elusive. This is partly caused by the lack of effort spent in exploring marine actinomycetes, whereas terrestrial actinomycetes have been, until recently, a successful source of novel bioactive metabolites. Furthermore, skepticism regarding the existence of indigenous populations of marine actinomycetes arises from the fact that the terrestrial bacteria produce resistant spores that are known to be transported from land into sea, where they can remain available but dormant for many years. Thus, it has been frequently assumed that actinomycetes isolated from marine samples are merely of terrestrial origin.

Table 1. List of marine actinomycetes reported so far from the Indian Peninsula

Sr. no.	Species	Habitat	Location
1	*Actinomycetes*sp.	Sediments	Managalavanam, Kerala
2	*Micromonospora*sp.	Sediments	Visakhapatnam coast, Andhra Pradesh(A.P.)
3	*Micropolyspora*sp.	Sediments	Visakhapatnam coast, A.P.
4	*Nocardia*sp.	Sediments	Visakhapatnam coast, A.P.
5.	*Rhodococcus*sp.	Oil polluted coastal region	Mumbai harbour, Mumbai
6	*Streptomyces albidoflavus*	Sediments	Pitchavaram mangrove, Tamil Nadu (T.N.)
7	*S. alboniger*	Sediments	Vellar estuary, T.N.
8	*S. albovinaceus*	Sediments	Vellar estuary, T.N.
9	*S. albus*	Different parts of fishes	Vellar estuary, T.N.
10	*S. aureocirculatus*.	Sediments	Pitchavaram mangrove, T.N.
11	*S. aureofasciculus*	Sediments	Vellar estuary, T.N.
12	*S. baarnensis*	Sediments	Vellar estuary, T.N.
13	*S. baarnensis*	Sediments	Vellar estuary, T.N.
14	*S. canus*	Different parts of fishes	Vellar estuary, T.N.
15	*S. chattanogensis*	Different parts of fishes	Vellar estuary, T.N.

16	*S. clavifer*	Sediments	Pitchavaram mangrove, T.N.
17	*S. fradiae*	Sediments	Arabian Sea
18	*S. galbus*	Alimentary canal of fishes	Vellar estuary, T.N.
19	*S. galtieri*	Sediments	Pitchavaram mangrove, T.N.
20	*S. gibsonii*	Sediments	Pitchavaram mangrove, T.N.
21	*S. griseobrunneus*.	Sediments	Vellar estuary, T.N.
22	*S. griseoflavus*	Sediments	Arabian Sea
23	*S. griseorubiginosus*	Sediments	Vellar estuary, T.N.
24	*S. hawaiiensis*	Different parts of fishes	Vellar estuary, T.N.
25	*S. kanamyceticus*.	Sediments	Pitchavaram mangrove, T.N.
26	*S. marinensis*	Water	Visakhapatnam coast, T.N.
27	*S. moderatus*	Sediments	Vellar estuary, T.N.
28	*S. nigrifaciens*	Sediments	Vellar estuary, T.N.
29	*S. olivoviridis*	Different parts of fishes	Vellar estuary, T.N.
30	*S. orientalis*	Different parts of fishes	Vellar estuary, T.N.
31	*S. palveraceus*	Sediments	Arabian Sea
32	*S. plicatus*	Alimentary canal of fish	Veli Lake, Kerala

33	*S. rimosus*	Gut contents of fish	Vellar estuary, T.N.
34	*S. roseolilacinus*	Sediments.	Pitchavaram mangrove, T.N.
35	*S. scabies*	Different parts of fishes	Vellar estuary, T.N.
36	*S. subflavus*	Sediments	Vellar estuary, T.N.
37	*S. vastus*	Sediments	Vellar estuary, T.N.
38	*S. violaceus*	Sediments	Vellar estuary, T.N.
39	*S. xantholiticus*	Sediments	Pitchavaram mangrove, T.N.
40	*Streptosporangium*sp.	Sediments	Visakhapatnam coast, A.P.
41	*Streptoverticillium*sp.	Sediments	Visakhapatnam coast, A.P.
42	*Saccharopolysporasalina*VITSDK4	Sediments	Marakkanam, Bay of Bengal, T.N.
43	*Streptomyces* VITSDK1 sp.	Sediments	Marakkanam, Bay of Bengal, T.N.
44	S. *fradiae* BW2-7	Sediments	Arabian Sea, Kerala

Source: Ranjith Kumar et al., 2014a; Sivakumar*et al*., 2007; Suthindhiran and Kannabiran, 2009a; Suthindhiran and Kannabiran, 2009b.

2.2 Streptomyces

The search for new antibiotics or new antibiotic producing microbial strains continues to be of utmost importance in research programs around the world because of the increase of resistant pathogens and toxicity of some used chemical antibiotics. Among actinomycetes a large number of antibiotics were obtained from the genus Streptomyces (Alan and James, 2007; Lyudmila *et al*., 2008; Junker *et al*., 2009; Koch and Loffler, 2009). *Streptomyces* are widely recognized

as industrially important microorganisms because of their ability to produce many kinds of novel secondary metabolites including antibiotics (Williams *et al*., 1983). Indeed, different Streptomyces species produce about 75% of commercially and medically useful antibiotics (Miyadoh, 1993).

Streptomyces species are distributed widely in marine and terrestrial habitats (Pathomaree*et al*., 2006) and are of commercial interest due to their unique capacity to produce novel metabolites. In fact, the genus Streptomyces alone accounts for a remarkable 80% of the actinomycete natural products reported to date, a biosynthetic capacity that remains without rival in the microbial world (Watve*et al*., 2001). It was also expected that Streptomyces species will have a cosmopolitan distribution, as they produce abundant spores that are readily dispersed (Antony-Babu*et al*., 2008). These filamentous bacteria are well adapted to the marine environment and are able to break down complex biological polymers. Marine Streptomyces are widely distributed in biological sources such as fishes, molluscs, sponges, seaweeds, mangroves, besides seawater and sediments. The genus Streptomyces was classified under the family Streptomycetaceae, which includes Gram-positive aerobic members of the order Actinomycetales and suborder Streptomycineae within the new class Actinomycetes (Stackebrandt*et al*., 1997; Anderson and Wellington, 2001). They have DNA G ± C content of 69–78 mol%. These organisms are gaining importance not only for their taxonomic and ecological perspectives, but also for their production of novel bioactive compounds like antibiotics, enzymes, enzyme inhibitors, pigments and for their biotechnological application such as probiotics and single cell protein.

3. Identification of actinomycetes

The traditional methods used for the identification of the aerobic filamentous actinomycetes are laborious, time consuming and often

require a series of specialized tests (Steingrube*et al*., 1995; Wilson *et al*., 1998; Harvey *et al*., 2001). Chemical criteria, such as the isomer of diaminopimelic acid (DAP) present in the cell wall and the diagnostic sugar(s) present in the whole-cell hydrolysate, have been used to separate the actinomycete genera into broad chemotaxonomic groups. However, determination of these characteristics is time-consuming and, in most cases, cannot identify an isolate to a single genus (Lechevalier, 1989). PCR-based methods have provided a rapid and accurate way to identify these bacteria (Gurtler*et al*., 1991; Kohler *et al*., 1991; Beyazova and Lechevalier, 1993; Telenti*et al*., 1993; Soini*et al*., 1994; Mehling*et al*., 1995; Steingrube*et al*., 1997; Wilson *et al*., 1998; Laurent *et al*., 1999).

By conventional isolation methods, members of the genus Streptomyces comprise more than 95% of the filamentous actinomycete population in soil (Lacey, 1973; Elander, 1987). The Streptomycetes produce more antibiotics than any other genus of bacteria and, therefore, have been heavily exploited as a source of novel antimicrobial agents (Watve*et al*., 2001). The probability of isolating known species of Streptomyces from the environment is thus great and the probability of isolating novel antibacterial molecules from such species is very low. The isolation of the rarer, non- Streptomyces actinomycetes greatly increases the probability of isolating novel antibacterial molecules (Lazzarini*et al*., 2000). Therefore, a rapid method to distinguish Streptomycetes from other actinomycetes and to identify the non Streptomycetes to the genus level would be extremely useful. This would be of particular value in discerning between Streptomycetes and non Streptomycetes, such as Actinomadura, Nocardia and Nocardiopsis, whose colonies may be morphologically similar on agar plates.

3.1. Molecular Approach

The most powerful approaches to taxonomy are through the study of nucleic acids. Because these are either direct gene products or the genes themselves and comparisons of nucleic acids yield considerable information about true relatedness. Molecular systematics, which includes both classification and identification, has its origin in the early nucleic acid hybridization studies, but has achieved a new status following the introduction of nucleic acid sequencing techniques (O'Donnell *et al.*, 1993). Significance of phylogenetic studies based on 16S rDNA sequences is increasing in the systematics of bacteria and actinomycetes (Yokota, 1997). Sequences of 16S ribosomal DNA have provided actinomycetologists with a phylogenetic tree that allows the investigation of evolution of actinomycetes and also provides the basis for identification. Analysis of the 16S rDNA begins by isolating DNA (Hapwood, 1985) and amplifying the gene coding for 16S rRNA using the polymerase chain reaction (*e.g.* SivaKumar, 2001). The purified DNA fragments are directly sequenced. The sequencing reactions are performed using DNA sequencer in order to determine the order in which the bases are arranged within the length of sample (Xu Li-Hua *et al.,* 1999) and a computer is then used for studying the sequence for identification using phylogenetic analysis procedures.

3.2. Chemotaxonomical Approach

Chemotaxonomy is the study of chemical variation in organisms and the use of chemical characters in the classification and identification. It is one of the valuable methods to identify the genera of actinomycetes. Studies of Cummins and Harris (1956) established that actinomycetes have a cell wall composition akin to that of gram-positive bacteria, and also indicated that the chemical composition of the cell wall might furnish practical methods of differentiating various types of actinomycetes. This is because of the fact that chemical components of

the organisms that satisfy the following conditions have significant meaning in systematics.

i. They should be distributed universally among the microorganisms studied; and,
ii. The components should be homologous among the strains

Within a taxon, while significant differences exist between the taxa to be differentiated. Presence of Diaminopimelic acid (DAP) isomers is one of the most important cell-wall properties of gram-positive bacteria and actinomycetes. Most bacteria have a characteristic wall envelope, composed of peptidoglycan. The 2, 6- Diaminopimelic Acid (DAP) is widely distributed as a key aminoacid and it has optical isomers. The systematic significance lies mostly in the key aminoacid with two amino bases, and determination of the key aminoacid is usually sufficient for characterisation. If DAP is present, bacteria generally contain one of the isomers, the LL-form or the *meso*-form, mostly located in the peptidoglycan.

The sugar composition often provides valuable information on the classification and identification of actinomycetes. Actinomycete cells contain some kinds of sugars, in addition to the glucosamine and muramic acid of peptidoglycan. The sugar pattern plays a key role in the identification of sporulatingactinomycetes which have meso-DAP in their cell walls. However, the actinomycetes which have LL-DAP along with glycine have no characteristic pattern of sugars (Lechevalier and Lechevalier, 1970) and hence the whole cell sugar test has not received much attention here.

3.3 Classical Approach

Classical approaches for classification make use of morpho logical, physiological, and biochemical characters. The classical method described in the identification key by Nonomura (1974) and Bergey's Manual of Determinative Bacteriology (Buchanan and Gibbons, 1974) is

very much useful in the identification of Streptomycetes. These characteristics have been commonly employed in taxonomy of Streptomycetes for many years. They are quite useful in routine identification. They are as follows, **Aerial Mass Color, Melanoid Pigments, Reverse Side Pigments, Soluble Pigments, Spore Chain Morphology, Spore Surface morphology, Assimilation of Carbon Source.**

3.4 Numerical Taxonomic Approach

Numerical taxonomy involves examining many strains for a large number of characters prior to assigning the test organism to a cluster based on shared features. The numerically defined taxa are polythetic, so no single property is either indispensable or sufficient to entitle an organism for membership of a group. Once classification has been achieved, cluster specific or predictive characters can be selected for identification (Williams *et al.*, 1983). Numerical taxonomy was first applied to *Streptomyces* by Silvestri*et al.* (1962). The numerical taxonomic study of the genus *Streptomyces* by Williams *et al.* (1983) involves determination of 139 unit characters for 394 type cultures of *Streptomyces.* Clusters were defined at 77.5% or 81% Ssm and 63% Sj similarity levels, and the former coefficient is being used to define the clusters. His study includes 23 major, 20 minor and 25 single member clusters.

The numerical classification of the genus *Streptomyces* by Kampfer*et al*. (1991) involves determination of 329 physiological tests. His study includes 15 major clusters, 34 minor clusters and 40 single member clusters which are defined at 81.5% similarity level Ssm using the simple matching coefficient (Sokal and Michener, 1958) and 59.6 to 64.6% similarity level Sj using Jaccard coefficient (Sneath, 1957). Thus, numerical taxonomy provides us with an invaluable framework for *Streptomyces* taxonomy, including identification of species.

4. Metabolite production by marine Streptomyces

Actinomycetes comprise about 10% of the bacteria colonizing marine aggregates and can be isolated from various marine sources. Many actinomycete isolates from the depths of the oceans contain non-ribosomal polyketide synthase (NRPS) and polyketide synthase (PKS) pathways, the hallmarks of secondary metabolite production (Li and Piel, 2002; Salmon *et al*., 2003). Terrestrial soils have hitherto been the predominant and widely exploited source, and investigations on marine Streptomyces are few and inconclusive, though they are the important sources for new bioactive compounds (Okami, 1984). About 23,000 antibiotics have been discovered from microorganisms. It has been estimated that approximately 10,000 of them were isolated from actinomycetes (Okami and Hotta, 1988). Actinomycetes, mainly the genus Streptomyces, have the ability to produce a wide variety of secondary metabolites as bioactive compounds, including antibiotics. The name Streptomyces was introduced in 1943 for the aerial mycelia producing actinomycetes. The genus Streptomyces is represented in nature by the largest number of species among all the genera of actinomycetes and figures over 500 species. The group has an enormous biosynthetic potential that remains unchallenged among other microbial groups. The immense diversity, along with itsunder utilization is the fundamental reason for attracting researchers towards it for discovering novel metabolites. During the last decade, there has been increasing number of novel metabolites possessing potent bioactivity isolated from marine-derived Streptomyces (Lam, 2006; Wu *et al*., 2007). Many of them are cytotoxic and come from a wide variety of chemical structures such as macrolides, a-pyrones, lactones, indoles, terpenes and quinones.

4.1. Antagonistic marine actinomycetes in Indian peninsula

Of 9 maritime states of India, only 4 have been extensively covered for the study of marine actinomycetes. Forty years of floristic

inventory of marine actinomycetes in Indian Peninsula yielded 41 species belonging to 8 genera, in which the genus Streptomyces was more frequently recorded(Sivakumar*et al*., 2007). Majority of the surveys have been conducted in the coastal areas, collecting the littoral sediments from the states of Maharastra, Kerala, Tamil Nadu and Andhra Pradesh. Studies covering the Gujarat, Goa, Karnataka, Orissa, West Bengal and Andaman and Nicobar islands are scanty.

Ranjithkumar et al., (2014a) isolated 115 actinomycetes from were isolated from the east and west coastal regions of India. Out of 115 actinomycetes strains, 52 isolates (45.2%) had antimicrobial activity, of which 26 isolates (22.6%) showed antibacterial activity (against *S. aureus*), 17 isolates (14.7%) showed antifungal activity (against *T. rubrum*), 9 isolates (7.8%) showed both antibacterial and antifungal activity. Of all 115 actinomycetes strains, BW2-7 showed a broad-spectrum antibiotic activity (Fig. 1) which was isolated from Kannamaly beach had broad spectral antimicrobial activity and was selected for further studies. The physiological and biochemical characteristics and the 16S rRNA sequence analysis confirmed that the strain *Streptomyces sp*BW2-7 was identical to *S. fradiae.*

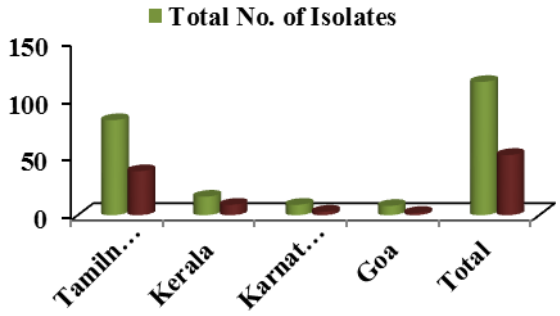


Fig.1: Antagonistic isolates from various regions of Indian Peninsula
Source: Ranjith Kumar et al., 2014

4.1.1 Tamil Nadu

Laksmanaperumalsamy *et al.* (1978) isolated 518 *Streptomyces* strains from the sediments of estuarine, backwater, marine, freshwater and mangrove environment of Porto Novo using Grein and Meyer's agar, Kuster's agar and Glucose asparagines agar. Majority of the isolates (46.43%) showed combined antibacterial and antifungal activity and 25% showed only antibacterial activity. Balagurunathan *et al.* (1989) studied the antagonistic behavior of actinomycetes isolated from the littoral sediments of Parangipettai. Among the 51 strains, only 11 strains showed good antibiotic activity and they were identified as *Streptomyces* spp. and *Nocardia*spp. This antibiotic was tested against fish pathogens *viz.* species of *Vibrio, Aeromonas, Pseudomonas, Bacillus* and *Fusarium*and it inhibited all these pathogenic organisms with inhibition zones ranging from 10-30 mm. The *Vibrio* sp. and *Pseudomonas* sp. were more sensitive than the other bacterial species tested.

Sivakumar (2001) isolated actinomycetes from the Pitchavaram mangrove environment. The 16S rRNA genes of the isolated two strains were partially sequenced were deposited in the Gen Bank, National Centre for Biotechnological Information, USA under the sequence of the accession numbers AY015427 and AY015428. Patil *et al.* (2001) reported 133 strains of actinomycetes from 129 marine samples collected from various stations along the Tuticorin coast. Of the 104 strains of actinomycetes screened for the inhibitory activity against bacterial pathogens associated with fish diseases (*Aeromonashydrophila, Aeromonassobria*and *Edwardsiellatarda*), 77 isolates possessed inhibitory activity to at least one of the pathogens. Balagurunathan and Subramanian (2001) isolated 51 strains of *Streptomyces* from the littoral sediments of Parangipettai coastal waters. Out of these, only eight strains showed very promising antibiotic activity against bacteria and fungi. These strains exhibited higher activity against gram-positive

bacteria than the gram-negative bacteria. Patil *et al.* (2001) isolated 20 actinomycetes strains from water and sediment samples of mangrove area of Tuticorin. The strains were checked for their antagonistic activity against seven shrimp bacterial pathogens. Among them, 83% showed good antagonistic activity against all the tested pathogens.

Sahu*et al.* (2004) isolated 40 strains of actinomycetes from the gut contents of three estuarine fishes*viz. Chanoschanos, Etroplussuratensis*and *Latescalcarifer*. Among them, only 10 strains of actinomycetes (30%) showed moderate antagonistic activity against all the tested bacterial pathogens. Kathiresan *et al.* (2005) isolated 160 strains from the sediments of mangrove, estuary, sand dune and industrially polluted marine environment of Cuddalore. Of the 160 isolates, 10 showed potent activity against all the fungi tested. These isolates produced high antifungal compounds at 120 h of incubation period in the production medium culture. Dhanasekaran *et al.* (2005) reported 107 strains of actinomycetes from 16 different marine soil samples and studied their antifungal activity against five test fungi. Out of these, only 22 isolate (21.2%) which were grown in starch casein agar produced diffusible antifungal substances in varying quantities. Potency of the culture filtrate was estimated by agar cup assay method using *C. albicans.* The antifungal activity was also tested by agar overlay method using *C. albicans*and *S. cerevisiae*as test organisms. Six isolates showed strong antifungal action in both agar cup and agar overlay assays.

Sivakumar *et al*. (2005a) reported 91 strains of actinomycetes from different stations of the Pitchavaram mangrove ecosystem. Out of the 91 strains, only 6 strains showed good activity and they were identified upto species level. Sahu *et al.* (2005a) studied actinomycetes population density from different samples *viz.* water, sediments, seaweeds, molluscs and finfishes of the Vellar estuary. The sediment samples harboured higher population density compared to the water

samples. Biological samples *viz.* seaweeds, molluscs and fin fishes were also analyzed for actinomycetes population. Among them, molluscs recorded higher population density in shell surface region than the gut contents, while the fin fishes recorded higher population in gut contents followed by gills and skin. Seaweed samples also recorded considerable actinomycetes populations. Sahu *et al.* (2005b) studied the extra-cellular enzyme (amylase, lipase, protease, cellulase and chitinase) activities of actinomycetes isolated from the sediment and molluscan samples of the Vellar estuary. The study indicated that the actinomycetes are the potential sources for extra-cellular enzymes, which play a role in biodegradation of organic matter, thereby enhancing the productivity of the marine environment.

Umamaheswary *et al.* (2005) isolated 40 strains of actinomycetes from the estuarine fish, *Mugilcephalus*using Kuster's agar medium. Out of 40 strains tested, only the strain *S. galbus*showed good L-glutaminase activity. Various process parameters which influenced L-glutaminase production by the *S. galbus*were optimized. Maximal enzyme production (18.93IU/ml) was attained at pH 9.0, 36^0 C, and glucose and malt-extract as carbon sources after 72 h of incubation. Senthilkumar *et al.* (2005) isolated 41 halophilic actinomycetes strains from the salt marsh area of the Vellar estuary using four different media. SC agar medium was the best for the isolation of halophilicactinomycetes. Among the isolated strains, the strain SH-9 showed greater resistance towards mercuric chloride in agar diffusion assay. The strain was classified as *Actinopolyspora*sp. by its morphological and chemotaxonomical characters. Sivakumar *et al.* (2006) isolated actinomycetes strains from skin, gills and gut contents of the estuarine fish, *Chanoschanos.* Out of 20 strains tested, *Streptomyces rimosus*showed L-glutaminase activity. Optimum production of L-glutaminase (18.93IU/ ml) was observed after 96 h at 27^0 C, pH 9.0 with glucose and malt extract.

Sahu *et al.* (2006) also reported a total number of 40 strains of actinomycetes from the sediments of the Vellar estuary and checked their antagonistic activity against the human bacterial pathogens (*B. substilis, P. vulgaris, S. flexineri, K. pneumoniae V. cholerae*and *S. aureus*). Among them, 9 strains (22.5%) showed activity against the tested pathogens and 5 strains which showed good activity were identified upto species level. Muthurayar *et al.* (2006) isolated a total of 18 actinomycetes strains from an estuarine fish, *Chanoschanos*and studied their antagonistic activity against human bacterial pathogens.

Saha *et al*. (2006) isolated four marine actinomycetes from Bay of Bengal and screened against multiple drug resistant (MDR) bacteria. Asha devi*et al*. (2006) isolated 3 marine actinomycetes from Dhanushkodi coastal region, Tamil Nadu. Vijayakumar *et al*. (2007) reported 192 actinomycetes colonies from 18 marine sediment samples of Palk Strait region of Bay of Bengal, India. Among them, 68 isolates were morphologically distinct on the basis of color of spore mass, reverse side color, aerial and substrate mycelia formation, production of diffusible pigment and sporophore morphology. From these isolates 39 were assigned to the genus Streptomyces. Gandhimathi *et al*. (2008) isolated 26 marine endosymbiotic strains, isolated from the Bay of Bengal. In this investigation Streptomyces species showed antimicrobial activity against pathogenic bacteria and fungi. Praveen *et al*. (2008) isolated two marine actinomycetes and optimized the fermentation conditions.

Selvin *et al*. (2009) optimized and produced antimicrobial agents from sponge associated marine actinomycetes*Nocardiopsisdassonvillei* MAD08. Suthindhiran and Kannabiran (2009a) isolated Streptomyces VITSDK1 spp from South coast of India which exhibited significant hemolytic activity against rat erythrocytes and human erythrocytes. It also showed moderate antibiosis against fungi and bacterial pathogens. Suthindhiran and Kannabiran (2009b) isolated *Saccharopoly*

*sporasalina*VITSDK4 from Bay of Bengal, which produced an extracellular bioactive metabolite, which inhibited the proliferation of HeLa cells as well as antagonistic to fungal and bacterial pathogens. Deepika and Kannabiran (2009) isolated 100 marine actinomycetes from coastal regions of Tamil Nadu. Out of 100 isolates 3 isolates belonged to genus *Streptomyces* sp exhibited potential antidermatophytic activity against the dermatophyte*Trichophytonrubrum*. Ramesh and Mathivanan (2009) isolated 209 marine actinomycetes from Bay of Bengal and screened for antimicrobial activity and industrial enzymes.

4.1.2. Kerala

Ranjithkumar et al. (2014b) isolated vancomycin from new source *S. fradiae*collected from kannamaly beach, Arabian sea. This vancomycin showed antiquorum sensing activity and wound healing properties again skin pathogens. Dhevendaran *et al*. (2004) isolated Streptomycetes from *Pernaviridis, Grapsusstrigosus, Ulva fasciata*and *Sargassumwightii*collected from Kovalam coast. The distribution pattern of the microorganisms with special emphasis on Streptomycetes was carried out using special microbiological media. Streptomycetes isolated from the visceral mass of *P. viridis*and *G. strigosus*showed maximum colonization in Actinomycete agar medium, whereas Streptomycetes associated with the fauna and seaweed showed a high diversity in pigmentation. *Streptomyces* harboured in the visceral mass of *P. viridis*exhibited antagonism against *Aeromonas*sp. Remya and Vijayakumar (2008) reported 173 actinomycetes from Kerala, West Coast of India. Out of these isolates 21 had antimicrobial activity.

4.1.3. Andhra Pradesh

Ellaiah and Reddy (1987) isolated 140 strains of actinomycetes from the marine sediments of Visakhapatnam coast and identified them upto genus level. Out of these 140 strains, only 18% exhibited anti-microbial activity against bacteria and fungi. Ellaiah (1996) isolated

actinomycetes from the sediments of Bay of Bengal and the strains which showed good antagonistic activity were identified upto species level. Ellaiah (2002) isolated 80 strains of actinomycetes from the sediments of Bay of Bengal near Machilipatnam by plating on starch casein agar medium. Of these, 7 isolates exhibited broad-spectrum antimicrobial activity, 68 showed proteolytic activity and 62 showed amylolytic activity. Ellaiah (2004) have isolated 60 actinomycetes from Bay of Bengal near Kakinada coast with distinct characteristics, by plating on starch casein agar medium. Among them, 11 isolates exhibited antibacterial (18.3%), 10 isolates showed antifungal (16.6%) while 2 isolates showed both antibacterial and antifungal (3.3%) activities. All 60 isolates were also tested for enzymatic activities on which 49 (81.6%) and 51 isolates (85%) exhibited amylolytic and proteolytic activities, respectively. Sujatha (2005) reported a new marine Streptomycetes BT-408 against methicillin resistant *S. aureus*.

4.1.4. Andaman and Nicobar group of islands

Sahu (2007) assessed the population density of actinomycetes from eight different stations of the Little Andaman Island. Mean population density of actinomycetes recorded from the water samples varied from 0.29 to 0.45 $x10^3$ CFU/ml with the minimum of 0.29 $x10^3$ CFU/ml at Navel Area and the maximum of 0.45 $x10^3$ CFU/ml at Chandra Nallah. In the case of sediment samples, population density ranged from 1.21 to 3.29$x10^3$ CFU/g with a minimum of 1.21 $x10^3$ CFU/g at Navel Area and a maximum of 3.29 $x10^3$ CFU/g at Buttler Bay. During the investigation, a total of 41 strains were isolated and tested for their antagonistic activity against the bacteria that are highly pathogenic to shrimps such as *V. alginolytics*, *V. harvey*iand *V. parahaemolyticus.* More than 61% of the strains (26 strains) exhibited varying degree of antagonistic activity. Among them, 6 strains showed good activity and they were tentatively identified. The results suggested that the

actinomycetes from the marine environment can be used as bio-control agents in shrimp culture systems to control diseases caused by bacterial pathogens.

5. Optimization of fermentation conditions

It is important to improve the performance of the systems and to increase the yield of the process without increasing the cost. The method used for this purpose is called optimization. There is a parameter change in the general practice of determining the optimal operating conditions while keeping the others at a constant level. This is called one-variable-at-a-time technique. The major disadvantage of this technique is that it does not depict the complete effects of the parameters on the process. In order to overcome this problem, optimization studies can be carried out using response surface methodology (RSM). To make the production of antibiotics feasible, it is necessary to optimize the production conditions. In statistical based approaches, response surface methodology (RSM) has been extensively used in fermentation media optimization (Dutta *et al*., 2004; Xiong*et al*., 2004).

5.1. Plackett- Burman (PB) experimental design

The purpose of the first step in the optimization strategy was to identify the medium components that have significant effect on the antibiotic production. Plackett–Burman designs are experimental designs presented in 1946 by Robin L. Plackett and J. P. Burman while working in the British Ministry of Supply. Their goal was to find experimental designs for investigating the dependence of some measured quantity on a number of independent variables (factors), each taking *L* levels, in such a way as to minimize the variance of the estimates of these dependencies using a limited number of experiments. Interactions between the factors were considered negligible. The solution to this problem is to find an experimental

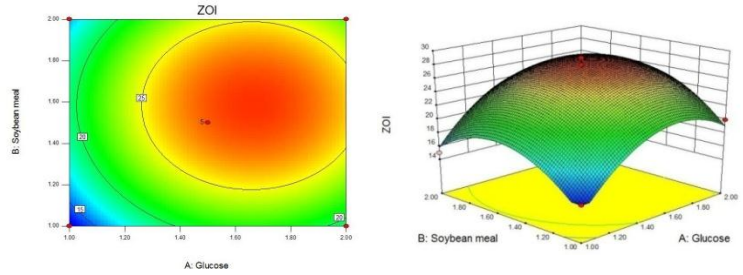
design where each combination of levels for any pair of factors appears the same number of times, throughout all the experimental runs. A complete factorial design would satisfy this criterion, but the idea was to find smaller designs. The PB design was based on the first-order model with no interaction among the factors.

5.2 Response Surface Methodology (RSM)

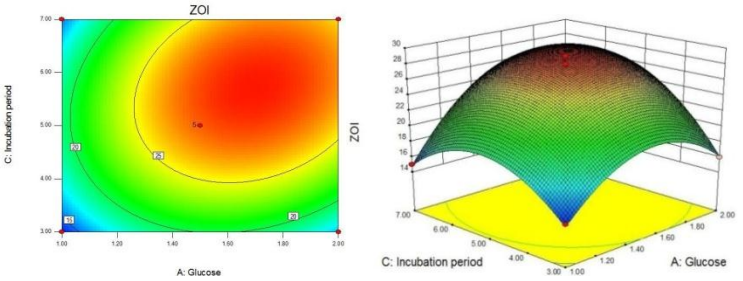
RSM has been very popular for optimization studies in recent years. It is clear that RSM has been widely applicable for different purposes in chemical and biochemical processes. RSM consists of a group of mathematical and statistical techniques that can be used to define the relationships between the response and the independent variables. RSM defines the effect of the independent variables, alone or in combination, on the processes. In addition to analyzing the effects of the independent variables, this experimental methodology also generates a mathematical model. The graphical perspective of the mathematical model has led to the term Response Surface Methodology (Myers and Montgomery, 1995; Anjum*et al*., 1997).

Response surface designs are commonly used to explore nonlinear relationships between independent (medium components) and the dependent (antimicrobial activity) variables (Rosenthal *et al*., 2001). Some computer packages offer optimal designs based on the special criteria and input from the user. These designs differ from one other with respect to their selection of experimental points, number of runs and blocks. After selection of the design, the model equation is defined and coefficients of the model equation are predicted. The visualization of the predicted model equation can be obtained by the response surface plot and contour plot. Ranjith Kumar et al. (2014a) reported that the Box-

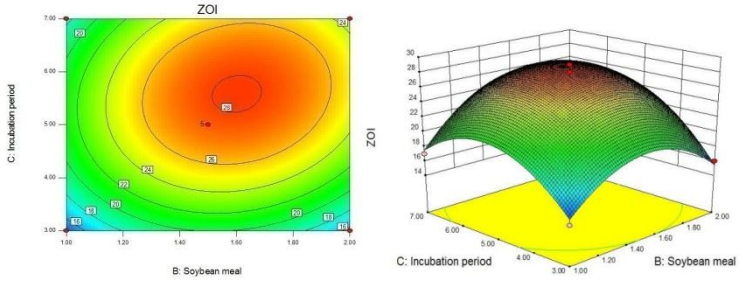
Fig. 2: Contour plots and 3-D plots for antibiotic production Zone of inhibition Vs Glucose and Soybean meal



Zone of inhibition Vs Glucose and Incubation period



Zone of inhibition Vs Soybean meal and Incubation period



Source: Ranjith Kumar et al., 2014

Behnken design was conducted in the optimum vicinity to locate the optimum concentration of Soybean meal, Glucose and Incubation period for maximum antibiotic production. These suggested that the

concentration of glucose and incubation period have a direct effect on antibiotic activity and hence yield. All of the above consideration indicated an excellent adequacy of the regression model (Fig. 2). This report proves that the production of antibiotic as secondary metabolites is profoundly influenced by the kind and quality of nutritional elements available and environmental factors.

References

1. Alan, T.B. and E.M.S. James. 2007. Marine actinobacteria: New opportunities for natural product search and discovery. *Trends Microbiology.* **15 (11**): 491-499.
2. Anderson, A.S. and E.M.H. Wellington. 2001. The taxonomy of Streptomyces and related genera. *International Journal of Systematic Evolutionary Bacteriology.* **51(3)**: 797-814.
3. Anjum, M.F., I. Tasadduq and K.Al-Sultan. 1997. Response surface methodology: A neural network approach. *European Journal of Operational Research*, **101**: 65-73.
4. Antony Babu, S., J.E.M. Stach and M. Goodfellow. 2008. Genetic and phenotypic evidence for *Staphylococcus griseus*ecovars isolated from a beach and dune sand system. *Antonie Van Leeuwenhoek*. **94(1)**: 63-74.
5. Asha devi, N.K., M. Jeyarani and K. Balakrishnan. 2006. Isolation and identification of marine actinomycetes and their potential in antimicrobial activity. *Pakistan Journal of Biological Sciences.* **9(3)**: 470-472.
6. Balagurunathan, R., G.S. Prasad, R. Manavalan and A. Subramanian. 1989. Actinomycetes from the littoral sediments of Parangipettai (South India) and their antibiotic activity. Proceedings of the First International Marine Biotechnology Conference- IMBC, Centre of Advanced study in Marine Biology, Annamalai University, Parangipettai, India. pp. 10
7. Baltz, R.H. 2006. Marcel Faber roundable: is our antibiotic pipeline improductive because of starvation, constipation or lack of inspiration? *Journal of Industrial Microbiology and Biotechnology.* **33(7)**: 507-513.
8. Baltz, R.H. 2008. Renaissance in antibacterial discovery from actinomycets. *Current Opinion in Pharmacology.* **8**: 557-563.

9. Berdy, J. 1995. Are actinomycetes exhausted as a source of secondary metabolites? Proceedings of the 9th International symposium on the biology of actinomycetes part1. *Allerton press*, New York. pp. 3-23.
10. Beyazova, M. and M.P. Lechevalier. 1993. Taxonomic utility of restriction endonuclease fingerprinting of large DNA fragments from Streptomyces strains. *International Journal of Systematic Bacteriology.* **43**: 674-682.
11. Buchanan, R.E. and N.E. Gibbons. 1974. Bergey's manual of Determinative Bacteriology. (Eighth edition), The Williams and Wilkins Co., Baltimore. pp. 747 - 842.
12. Challis, G.L. 2008. Mining Microbial genomes for new natural products and biosynthetic pathways. *Microbiology.* **154**: 1555-1569.
13. Cummins, C.S. and H. Harris. 1956. A Comparison of cell wall composition in *Nocardia*, *Actinomyces*, *Mycobacterium* and *Propion bacterium*. *Journal of GeneralMicrobiology*. **15**: 9.
14. Deepika, T.L. and K. Kannabiran. 2009. A report on antidermatophytic activity of actinomycetes isolated from Ennore coast of Chennai, Tamil Nadu, India. *International Journal of Integrative Biology*. **6**: 132-136.
15. Dhanasekaran, D., A. Panneerselvam and N. Thajuddin. 2005. Antibacterial actinomycetes in marine soils of Tamil nadu. *Geobios*. **32**: 37-40.
16. Dhevendaran K and P.K. Praseetha. 2004. Studies on streptomycetes associated with seaweed of Cape-Comarin, Tamilnadu. *Seaweed Research and Utilization.***26**:245-252
17. DiMasi, J.A., M.A. Seibring and L. Lasagna. 1994. New drug development in the United States from 1963 to 1992. *Clinical Pharmacology Therapy*. **55**: 609-622.

18. Dutta, J.R., P.K. Dutta and R. Banerjee. 2004. Optimization of culture parameters for extracellular protease production from a newly isolated *Pseudomonas* sp. using response surface and artificial neural network models. *Process Biochemistry.***39**: 2193-2198.
19. Elander, R.P. 1987. Microbial screening, selection and strain improvement. In Basic Biotechnology. Edited by Bulock J and Kristiansen B. London. *Academic press.*pp. 217-251.
20. Ellaiah, P. and A.P.C. Reddy. 1987. Isolation of actinomycetes from marine sediments of Visakhapatnam, east coast of India. *IndianJournal of marine Sciences.***16**: 134-135.
21. Ellaiah, P., D. Kalyan, V.S.V. Rao and B.V.L.N. Rao. 1996. Isolation and characterization of bioactive actinomycetes from sediments. *Hindustan Antibiotics Bulletin.***38**: 48-52
22. Ellaiah, P., K. Adinarayana, A. Naveen Babu, B. Thaer, T. Srinivasulu and T. Prabhakar. 2002. Bioactive acyinomycetesfrommarine sediments of Bay of Bengal near Machilipatnam. *Geobios.***29**: 97-100.
23. Ellaiah, P., T. Ramana, K.V.V.S.N. Bapi Raju, P. Sujatha and A. UmaSankar. 2004. Investigation on marine actinomycetes from Bay of Bengal near Kakinada coast of Andhra Pradesh. *AsianJournal of Microbiology, Biotechnology and Environmental Sciences.***6**: 53-56.
24. Fenical W. and P.R. Jensen. 1993. In: D.H. Attaway and O.R. Zaborsky (eds) Marine biotechnology I: Pharmaceutical and bioactive natural products. Plenum, New York.
25. Gandhimathi, R., M. Arunkumar, J. Selvin, T. Thangavelu, S. Sivaramakrishnan, G.S. Kiran, S. Shanmughapriya and K. Natarajaseenivasan. 2008. Antimicrobial potential of sponge

associated marine actinomycetes. *Journal of Medical Mycology*.**18**: 16-22.

26. Good fellow, M. and E. Williams. 1983. Ecology of actinomycetes. *Annual Review of Microbiology*. **37**: 189-216.
27. Good fellow, M. and J.A. Haynes. 1984. Actinomycetes in marine sediments.
    In Biological, Biochemical and Biomedical aspects of Actinomycetes, edited by Ortiz LO, Bojalil LF and Yakoleff V. *Academic press*. *Inc, Orlando Fl*. pp. 453.
28. Gurtler, V., V.A. Wilson and B.E. Mayall. 1991. Classification of medically important clostridia using restriction endonuclease site differences of PCR-amplified 16s rDNA. *Journal of General Microbiology*. **137**: 2673-2679.
29. Hapwood, D.A., M.J. Bill, K.F. Charter, T. Kieser, C.J. Bruton, H.M. Kieser, D. J. Lydiate, C.P. Smith, J.M. Ward and H. Schrempf. 1985. Genetic manipulation of *Streptomycetes*: A laboratory manual, John Innes Foundation, Norwich, *United Kingdom*. pp. 71-80.
30. Harvey, I., Y. Cormier, C. Beaulieu, V.N. Akimov, A. Meriaux and C. Duchaine. 2001. Random amplified ribosomal DNA restriction analysis for rapid identification of thermophilicactinomycete-like bacteria involved in hyper sensitivity pneumonitis. *Systematic and Applied Microbiology*. **24**: 277-284.
31. Helme, E. and H. Weyland. 1984. *Rhodococcusmarinonascens* sp. Nov., an actinomycete from the sea. *International Journal of Systematic Bacteriology*.**34**: 127-138.
32. Jensen, P.R., R.D. Wight and W. Fenical. 1991. Distribution of actinomycetes in near-shore tropical marine sediments. *Applied and Environmental Microbiology*. **57**: 1102-1108.

33. Junker, B., A. Walker, M. Hesse, M. Lester, J. Christensen and N. Connors. 2009. Actinomycetes scale- up for the production of the antibacterial, nocathiacin. *Biotechnology Progress*. **25(1)**: 176 -188.
34. Kampfer, P., R.M. Kroppenstedt and W. Dott. 1991. A numerical classification of the genera *Streptomyces* and *Streptoverticillium*using miniaturized physiological tests. *Journal of General Microbiology*. **137**: 1831-1891.
35. Kathiresan, K., R. Balagurunathan and M. MasilamaniSelvam. 2005. Fungicidal activity of marine actinomycetes against phytopathogen ic fungi. *Indian Journal of Biotechnology*. **4**: 271-276.
36. Kitagawa, W. and T. Tamura. 2008. Three types of antibiotics produced from *Rhodococcuserythropolis* strains. *Microbes and Environments.***23**: 163-171.
37. Koch, E. and L. Loffler. 2009. Partial characterization of the antimicrobial activity of *Streptomyces antimycotieus*. FZB53. *Journal of Phytopathology.***157 (4)**: 235-242.
38. Kohler, G., W. Ludwig and K.H. Schleifer. 1991. Differentiation of Lactococci by rRNA gene restriction analysis. *FEMS Microbiology Letters.***84**: 307-312.
39. Lacey, J. 1973. Actinomycetes in soils composts and fodders. In actinomycetes characteristics and practical importance. Edited by Sykes, G and Skinner,F.A. London. *Academic press.*pp. 231-251.
40. Laksmanaperumalsamy, P., D. Chandramohan and R. Natarajan. 1978. Antibacterial and antifungal activity of *Streptomycetes* from Porto Novo coastal environment. *Marine Biology.***1**: 15-24.
41. Lam, K.S. 2006. Discovery of novel metabolites from marine actinomycetes. *Current Opinion in Microbiology*. **9**: 245-251.
42. Larkin, M.J., L.A. Kulakov and C.C.R. Allen. 2005. Biodegradation and Rhodococcus masters of catabolic versatility. *Current opinion* inBiotechnology.**16**: 282-290.

43. Laurent, F.J., F. Provost and P. Boiron. 1999. Rapid identification of clinically relevant *Nocardia* species to genus level by 16S rRNA gene PCR. *Journal of Clinical Microbiology*. **37**: 99-102.
44. Lazzarini, A., L. Cavaletti, G.Toppo and F. Marinelli. 2000. Rare genera of actinomycetes as potential producers of new antibiotics. *Antonie van Leeuwenhoek*. **78**: 399-405.
45. Lechevalier, H.A. 1989. A practical guide to generic identification of actinomycetes. *In Bergey's manual of systematic Bacteriology*.. Edited by, Williams ST Sharpe ME and Holt JG. Baltimore: Williams and Wilkins. **4:** 2344-2347.
46. Lechevalier, M.P. and H.A. Lechevalier. 1970. Chemical composition as a criterion in the classification of aerobic actinomycetes. *International Journal of Systematic Bacteriology*. **20**: 435-443.
47. Li, A. and J. Piel. 2002. A gene cluster from a marine Streptomyces encoding the biosynthesis of the aromatic spiroketalpolyketideGris eorhodin A*. Journal of Chemical Biology*. **9**: 1017-1026
48. Lyudmila, A.R., U. Masataka, I.K. Natalia and V.M. Valery. 2008. Isolation, phylogenetic analysis and screening of marine mollusc-associated bacteria for antimicrobial, hemolytic and surface activities. *Microbiological Research.***163 (6)**: 633-644.
49. Mehling, A., U.F.Wehmeier and W. Piepersberg. 1995. Nucleotide sequences of Streptomycetes 16s ribosomal DNA toward a specific identification system for Streptomyces using PCR. *Microbiology*. **141**: 2139-2147.
50. Miyadoh, S. 1993. Research on antibiotic screening in Japan over the last decade a producing microorganisms approach. *Actinomycetologica*. **9**: 100-106.
51. Muthurayar, T., K. Sivakumar, M.K. Sahu, T. Thangaradjou and L. Kannan. 2006. Mutational effect on the antibacterial activity of

marine actinomycetes isolated from *Chanoschanos,* (Forskal, 1775). *Environment and Ecology.* **24**: 46-50.

52. Myers, R.H. and D.C. Montgomery. 1995. Response surface methodology: Process and product optimization using designed experiments. New York: John Wiley and Sons, Inc.
53. Nonomura, H. 1974. Key for classification and identification of 458 species of the *Streptomycetes*included in ISP. *Journal of Fermentation Technology*. **52(2)**: 78-92.
54. O'Donnell, A.G., T.M. Embley and M. Goodfellow. 1993. Future of Bacterial Systematics. In: Handbook of New Bacterial Systematics, London: *Academic Press*. pp.513-524.
55. Okami Y. 1984. Marine microorganisms as a source of bioactive agents. In: Klug J, Reddy CA (eds) Current perspective in microbial ecology. *American Society for Microbiology, Washington*. pp. 615-655.
56. Okami, Y. and K. Hotta. 1988. Search and discovery of new antibiotics. In: Goodfellow, M, Williams, S.T, Mordarski, M. (eds). Actinomycetes in biotechnology. *Academic press, London*. pp. 33-67.
57. Patil, R., G. Jeyaskaran and S.A. Shanmugan. 2001. Occurrence and activity of marine actinomycetes against shrimp bacterial pathogens. *Appl Fish Aqua.***1**: 79-81.
58. Plackett, R.L. and J.P. Burman. 1946. The design of optimum multifactorial experiments. *Biometrika.* ***33*****(4)**: 305-325.
59. Praveen, V., C.K.M. Tripathi, V. Bihari and S.C. Srivastava. 2008. Production of actinomycin-Dn by the mutant of a new isolate of *Streptomyces sindenensis*. *Brazilian Journal of Microbiology*. **39**: 689-692.
60. Ramesh, S. and N. Mathivanan. 2009. Screening of marine actinomycetes isolated from the Bay of Bengal, India for antimicrobial

activity and industrial enzymes. *World Journal of Microbiology and Biotechnology*. **25(12)**: 2103-2111.

61. Antibiofilm Formation and Wound Healing Activity of Glycopeptide Antibiotic Vancomycin Against Skin Pathogens. *International journal of Pharma and Biosciences*. **5(1):** 232 – 248.
62. Ranjith Kumar M,BrindhaPriyadarisini V, Srigopal Ram S, Senthil Kumar Thamilarasan, Ill Sup Nou.2014. Studies on a marine *Streptomyces fradiae* BW2-7 producing glycopeptide antibiotic Vancomycin effective against skin pathogens.Scholars Academic Journal of Biosciences. 2(11):746-761.
63. Ravel, J., M.J. Amoroso, R.R. Colwell and R.T. Hill. 1998. Mercury-resistant actinomycetes from the Chesapeake Bay. *FEMS Microbiology Letters*. **162**: 177-184.
64. Remya, M. and R. Vijayakumar. 2008. Isolation and characterization of marine antagonistic actinomycetes from west coast of India. *Factauviversitatis.***15**: 13-19.
65. Roessner, C.A. and Scott, A.I. 1996. Genetically engineered synthesis of natural products: from alkaloids to corrins. *Annual Review of Microbiology.***50**: 467-490.
66. Rosenthal, A., D.L. Pyle, K. Niranjan, S. Gilmour and L. Trinca. 2001. Combined effect of operational variables and enzyme activity on aqueous enzymatic extraction of oil and protein from soybean. *Enzyme and Microbial Technology*. **28**: 499-509.
67. Sahu, M.K., K. Sivakumar and L. Kannan. 2004. Estuarine fish as a source of antagonistic actinomycetes: An inventory.In: Proc Con MicrobioTropi Sea NIO, Goa, MB (O). pp. 04.
68. Sahu, M.K., K. Sivakumar and L. Kannan. 2005a. Isolation of actinomycetes from different samples of the Vellar estuary, South east coast of India. *Poll Res.* **24**: 45-48.

69. Sahu, M.K., K. Sivakumar and L. Kannan. 2005b. Degradationof organic matters by the extra-cellular enzymes of actinomycetes isolated from the sediments and molluscsof the Vellar estuary. *Journal of Aquatic Biology.***20 (2)**:142-144.
70. Sahu, M.K., K. Sivakumar and L. Kannan. 2006. Isolation and characterization actinomycetes inhibitory to human pathogens. *Geobios.***33 (2-3)**:105-109.
71. Sahu, M.K., K. Sivakumar, T. Thangaradjou and L. Kannan. 2007. Phosphate solubilizing actinomycetes in the estuarine environment: An inventory. *Journal of Environmental Biology*. **28(4):** 795-798.
72. Salmon, C.E., N.A. Magarvey and D.H. Sherman. 2003. Merging the potential of microbial genetics with biological and chemical diversity: an even brighter future for marine natural product drug discovery. *Nature product Report*. **21**: 105-121.
73. Selvin, J., S. Shanmugapriya, R. Gandhimathi, G.S. Kiran, T.R. Ravji, K. Natarajaseenivasan and T.A. Hema. 2009. Optimization and production of novel antimicrobial agents from sponge associated marine actinomycetes*Nocardiopsisdassonvelli*MAD08. *Applied Microbiology and Biotechnology*.
74. Senthilkumar, S., K. Sivakumar and L. Kannan. 2005. Mercury resistant halophilicactinomycetes from the salt marsh environment of the Vellar estuary, South east coast of India. *Journal of Aquatic Biology.* **20 (1):** 141-145.
75. Silvestri, L.G., M. Turri, L.R. Hill and E. Gilardi. 1962. A quantitative approach to the systematics of actinomycetes based on overall similarity. *ProceedingsofSymposium of the Society for General Microbiology*. **12**: 333- 360.

76. Sivakumar, K. 2001. Actinomycetes of an Indian mangrove (Pitchavaram) environment; an inventory Ph.D. Thesis, Annamalai University, Tamil Nadu.
77. Sivakumar, K., M.K. Sahu and K. Kathiresan. 2005a. Isolation and characterization of Streptomycetes, producing antibiotic from a mangrove environment. *Asian Journal of Microbiology, Biotechnology and Environmental Sciences.***7:** 87-94.
78. Sivakumar, K., M.K. Sahu, P.R. Manivel and L. Kannan. 2006. Studies on L-glutaminase producing actinomycetes strainLG-10 from the estuarine fish, *Chanoschanos* (Forskal, 1775). *Indian Journal of Experimental Biology*. **44**: 256-258
79. Sivakumar, K., M.K. Sahu, T. Thangaradjou and L. Kannan. 2007**.** Research on Marine actinobacteria in India, *Indian Journal of Microbiology*. **47**: 186-196.
80. Sneath, P.H.A. 1957. The application of computers to taxonomy. *Journal of General Microbiology.***17**: 201-226.
81. Soini, H., E.C. Bottger and M.K. Viljanen. 1994. Identification of mycobacteria by PCR- based sequence determination of the 32-Kilodalton protein gene. *Journal of Clinical Microbiology.***32**: 2944-2947.
82. Sokal, R. and C.D. Michener. 1958. A statistical method for evaluating systematic relationship. *Kanas University Science Bulletion*. **38**: 1409 -1438.
83. Srinivasan, M.C., R.S. Laxman and M.V. DeshPande. 1991. Physiology and nutrition aspects of actinomycets- An overview. *World Journal of Micobiology and Biotechnology***7**: 171-184.
84. Stackebrandt, E., F.A. Rainey and N.L. Ward Raine. 1997. Proposal for a new hierarchic classification system, Actinobacteria classis nov.*International Journal of Systematic Bacteriology*. **47**: 479-491.

85. Steingrube, V.A., B.A. Brown, J.L. Gibson and 7 other authors. 1995. DNA amplification and restriction endonuclease analysis for differentiation of 12 species and taxa of *Nocardia*, including recognition of four new taxa within the *Nocardia* asteroids complex. *Journal of Clinical Microbiology*. **33**: 3096-3101.

86. Steingrube, V.A., R.W. Wilson, B.A. Brown, K.C. Jost, Z.J.R. Blacklock, J.C. Gibson and R.J.J.R. Wallace. 1997. Rapid identification of the clinically significant species and taxa of aerobic actinomycetes, including *Actinomadura, Gordona, Nocardia, Rhodococcus, Streptomyces* and*Tsukamarella* isolated by DNA amplification and restriction endonuclease analysis. *Journal of Clinical Microbiology*. **35**: 817-822.

87. Strohl, W.R. 1997. Industrial antibiotics; today and the future. In Strohl WR (edited) *Biotechnology of antibiotics*, 2nd edn. Marcel Oekker, New York. pp. 1-47.

88. Sujatha, P., K.V. Bapiraju and T. Ramana. 2005. Studies on a new marine Streptomycete BT 408 producing polyketide antibiotic SBR-22 effective against methicillin resistant *Staphylococcus aureus*. *Microbiological Research*. **60**: 119-126.

89. Suthindhiran, K. and K. Kannabiran. 2009a. Cytotoxic and antimicrobial potential of actinomycete species *Saccharopolysporasal ina*VITSDK4 isolated from the Bay of Bengal Coast of India. *American Journal of Infectious Diseases*. **5**: 90-98.

90. Suthindhiran, K. and K. Kannabiran. 2009b. Hemolytic activity of *Streptomyces* VITSDK1 spp. Isolated from marine sediments in Southern India. *Journal of Medical Mycology*. **9**: 77-86.

91. Takizawa, M., R.R. Colwell and R.T. Hill. 1993. Isolation and diversity of actinomycetes in the Chesapeake Bay. *Applied and Environmental Microbiology*. **59**: 997-1002.

92. Telenti, A., F. Marchesi, M. Balz, F. Bally, E.C. Bottger and T. Bodmer. 1993. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. *Journal of Clinical Microbiology.* **31**: 175-178.
93. Umamaheswary, K., M.K. Sahu, K. Sivakumar, T. Thangaradjou, D. Sumitha and L. Kannan. 2005. Investigations on L-glutaminase producing actinomycetes strain LG-33 from the estuarine fish, *Mugilcephalus* (Linnaeus, 1758). *Environment and Ecology.***23**: 942-947.
94. Umezawa, H. 1982. Low-molecular-weight inhibitors of microbial origin. *Annual Review of Microbiology.***36**: 75-99.
95. Ventura. M., M. Canchaya, A. Tauch, G. Chandra, G.F. Fitzgerald, K.F. Chater and D. Van Sinderen. 2007. Genomics of actinobacteria: Tracing the evolutionary history of an ancient phylum. *Microbiology and Molecular Biology Reviews*. **71**: 495-548.
96. Vijayakumar, R., C. Muthukumar, N. Thajuddin, A. Pannerselvam and R. Saravanamuthu. 2007. Studies on the diversity of Actinomycetes in the Palk Strait region of Bay of Bengal, India. *Actinomycetologica.* **21**: 59-65.
97. Wenzel, S.C. and R. Muller. 2009. The impact of genomics on the exploitation of the myxobacterial secondary metabolome. *Nature product Report.***26**: 1385-1407.
98. Weyland, H. 1981. Distribution of actinomycetes on the sea floor. Zentrabl. Bankeriol. Parasitenkd. Infektionskr. Hy.G. Abt.1 Suppl. **11**: 185-193.
99. Weyland, H. and E. Helmke. 1988. Actinomycetes in the marine environment. In the biology of actinomycetes. Proceedings of the Biology of Actinomycetes, ed by Okami Y, Beppu J and Ogamura H. *Japan Scientific Society Press, Tokyo*. 294.

100. Williams, S.T., M. Goodfellow, G. Alderson, E.M. Willington, P.H. Sneath and M.J. Sacki. 1983. Numerical classification of Streptomyces and related genera. *Journal of General Microbiology.***129**: 1747-1813.
101. Wilson, R.W., V.A. Steing rube, B.A. Brown and R.J.J.R. Wallace. 1998. Clinical application of PCR- restriction enzyme pattern analysis for rapid identification of aerobic actinomycete isolates. *Journal of Clinical Microbiology*. **36**: 148-152.
102. Wu, S.J., S. Fotso, F. Li, S. Qin and H Laatsch. 2007. Amorphanesesquiterpenes from a marine Streptomyces sp. *Journal of Natural Product.***70**: 304-306
103. Xiong, Y.H., J.Z. Liu, H.Y. Song and L.N. Ji. 2004. Enhanced production of extracellular ribonuclease from *Aspergillusniger*by optimization of culture conditions using response surface methodology. *Biochemical Engineering Journal*. **21**: 27-32.
104. Xu, L.H., X. Jin, P.M. Mao, Z.F. Lu, X.L. Cui and C.L. Jiang. 1999. Three new species of the genus Actinobispora of the family Pseudonocardiaceae, *Actinobisporaalaniniphila*sp. nov., *Actinobisporaaurantiaca*sp. Nov and *Actinobisporaxinjiangensis* sp. Nov. *International Journal of Systematic Bacteriology.* **49**: 881-886.
105. Yokota, A. 1997. Phylogenetic relationship of actinomycetes. Atlas of actinomycetes, Asakura Publishing Co. Ltd., Japan. pp. 194 - 197.

Chapter 10
Biotechnology in Forensic Sciences

Harendra Modak and Madhuri Biradar

1. History of Forensic Genetic Typing

More than two decades ago, genetically, polymorphic protein markers were applied to potentially distinguish individuals. But, numerous factors restricted the forensic use of these protein-based genetic systems. The refining power of these markers was low, resultantly, individualization was not difficult. Additionally, the proteins are not available at satisfactory levels for typing in most of the tissues, and they are comparatively more prone to degradation in biological samples in contact with environmental fluctuations.

The typing of genetic polymorphisms at the DNA level assists to overcome these restrictions to a much greater extent. Primarily, there is a remarkable amount of dissimilarities at the DNA level to exploit for individuality testing. Second, any biological substance that comprises nucleated cells potentially can be used for DNA polymorphism typing. Third, DNA is found to be more stable in forensic samples. Consequently, with the current techniques for DNA typing and the set of existing genetic markers, human polymorphism typing at the DNA level is more delicate, more precise, and more instructive than the conventional protein genetic markers. Additionally, DNA technology provides the forensic scientist the greatest possibility to ignore individuals who have

been falsely connected with a biological sample and to decrease the number of individuals hypothetically incorporated as contributors of the sample to a limited (if not one) individuals.

The early method used regularly for human identity testing (from the mid-1980s) in the US and many other countries was restriction fragment length polymorphism (RFLP) typing of VNTR (variable number of tandem repeat) loci. Typing VNTR loci by RFLP investigation delivered a very high power of judgement, but required samples should hold minimum 10–25 ng of DNA that need to be some what integrated with fragment lengths up to 10,000 bp for any successful typing. Gradually, RFLP analysis was substituted by amplification of DNA (Fig.1) loci by the polymerase chain reaction (PCR) followed by typing of genetic markers. PCR facilitates the scrutinization of samples containing pictogram intensities of DNA, thus enhancing the preciseness of detection one to two folds of magnitude. In fact, small quantities of DNA extracted from different types of samples are typed customarily and effectively using PCR-based assays in forensic laboratories. Sample materials comprise saliva, blood, sweat and semen dropped on different substrates including cigarettes, clothing, postage stamps, drinking straws and containers, envelope flaps, face masks and chewing gum; various tissues from human remains; vaginal swabs from a rape victim; and possible reference samples taken from personal items, such as tooth brushes, hair brushes and razors, which may be advantageous in the identification of anonymous remains.

The initial genetic marker systems examined using PCR-based systems were centred on SNP variation. These SNP-centred systems offered high sensitivity of exposure but did not deliver the power of discernment that VNTR/RFLP typing afforded, and were not very useful for revealing the role of mixed samples because of their limited polymorphism.

To intensify the influence of discrimination of PCR-centred systems, VNTR loci were amplified by polymerase chain reaction, with the allelic forms fragmented by electrophoresis, and successively perceived by silver staining. Short tandem repeat (STR) or microsatellite loci, part of a subclass of highly polymorphic VNTR loci, have substituted the PCR-centred genetic markers. Nowadays, they are used in forensic laboratories world wide sincethey have polymorphic nature, small product size and semi-automated analytical methods. These repeat sequence loci are abundant in the human genome and are highly polymorphic. Autosomal STR loci are comprised of tandemly repeated sequences, each of which is made up of only four to five BP in length. The variable number of alleles of the forensically applicable STR loci ranges usually from 5 to 20 common alleles. The regions covering the repeats are normally small in size and are therefore responsive to amplification by PCR. Because the product length is generally less than 350 bp, much smaller than the length of a DNA fragment prerequisite for VNTR analysis by RFLP, slightly degraded samples are currently typeable. To lessen sample consumption and reduce laboratory manipulations, the STR loci are studied in a multiplex fashion, with approx 15 STR loci amplified and typed at the same time.

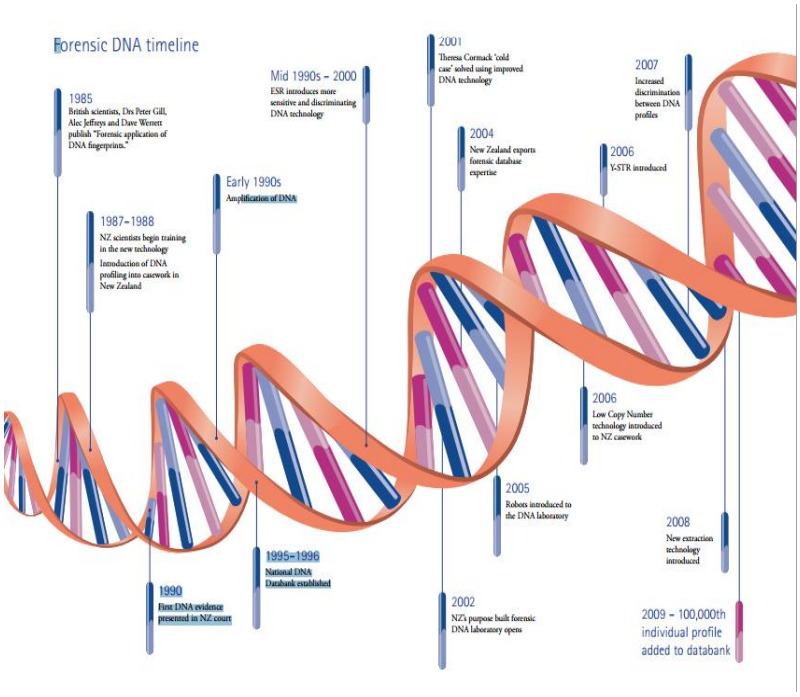

Forensic DNA timeline
1985
British scientists, Drs Peter Gill, Alec Jeffreys and Dave Werrett publish "Forensic application of DNA fingerprints."
1987–1988
NZ scientists begin training in the new technology
Introduction of DNA profiling into casework in New Zealand
Early 1990s
Amplification of DNA
1990
First DNA evidence presented in NZ court
1995–1996
National DNA Databank established
Mid 1990s – 2000
ESR introduces more sensitive and discriminating DNA technology
2001
Theresa Cormack 'cold case' solved using improved DNA technology
2002
NZ's purpose built forensic DNA laboratory opens
2004
New Zealand exports forensic database expertise
2005
Robots introduced to the DNA laboratory
2006
Low Copy Number technology introduced to NZ casework
2006
Y-STR introduced
2007
Increased discrimination between DNA profiles
2008
New extraction technology introduced
2009 – 100,000th individual profile added to databank

Figure 1: History of Biotechnological advancement in Forensic Sciences.

The detailed history of Forensic science is summarised hereunder:

- 1686 –Anatomy Professor Marcello Malpighi mentions in his treaties the ridges, spirals and loops in fingerprints.
- 1839 – H. Baynardprints the first reliable techniques for the detection of sperm, and examines the various microscopic characteristics of various different substrate fabrics.
- 1879 –Rudolph Virchow, a German pathologist, is one of the first to study about hair
- 1892 – Sir Francis Galton in his book "Fingerprints" establishes the individuality of fingerprints and a first classification system for it.

- 1892 – Juan Vucetich develops the fingerprint classification system that was used in Latin America.
- 1901 – Karl Landsteiner notices human blood group.
- 1915 – Prof. Leone Lattes develops the antibody test for ABO blood types.
- 1931 – Franz Josef Holzercomes up with the absorption-inhibition AB typing techniques that were generally used in forensic laboratories.
- 1940 – Rh blood groups were first described by Landsteiner and A. S. Weiner.
- 1953 – James Watson and Francis Crick identify the structure of DNA.
  1984 – Sir Alec Jeffreys discovers a method of identifying individuals by DNA –RFLP.
- 1986 – Kerry Mullis discovers PCR method.
- 1986 – DNA is used for the first time to solve a crime by Alec Jeffreys.
- 1987 – DNA profiling is introduced for the first time in a US criminal court.
- 1992 – Thomas Caskey and colleagues suggest the use of short tandem repeats for forensic DNA analysis in their published paper.
- 1996 – Mitochondria DNA evidence is used in US court for the first time.
- 1998 – NIDIS, FBI DNA database, is put into practice.

2. Introduction

"It has long been an axiom of mine, that the little things are infinitely the most important."

----------Arthur Conan Doyle

It was never been thought that the DNA molecule, a little thing, could prove to be possibly the most important source in the crime investigation. Since the onset of DNA fingerprinting, scrutinization of DNA molecules helps in identifying victims of crimes or accidents and

convicts or exonerates the suspects. It inspires the development of new methods in molecular biology, statistical analysis and usage of databases.DNA evidence speaks more than an eye witness.

The word forensic comes from the Latin word forensic which means' "*of or before the forum*". The usage of advanced scientific techniques for the implementation of law during criminal and civil cases to get the solution of important questions about the crime is known as forensic science. It includes both science and law.

Forensic investigation is the use of the tools of science along with exact scientific facts, to help in solving legal problems. All the branches of forensic investigation don't include biotechnology. For example, the study of fingerprinting or firearms proof does not include biotech. Nevertheless, examination of proof from blood and bodily fluids does depend on biotechnology.

Several types of investigators are applied in these analyses. Crime scene investigators regulate access to the scene to evade any external contamination. Then, they collect evidence for laboratory assessment. The samples must be picked out by avoiding cross-contamination with other evidence on the crime scene. The supervision of this evidence must be closely valued for any typesof poor storage, contamination or tampering. The evidence is then cautiously assessed in a laboratory. Laboratory crime scene investigators evaluate tissues or blood evidence and perform many other experiments to examine the samples collected at the scene.

Forensic methods to categorize someone have progressed from considering a person's actual fingerprints (looking at the arches and whorls in the fingertips) to analyzing genetic fingerprints. DNA fingerprinting is also called DNA profiling or DNA typing. Though human DNA is 99% - 99.9% identical from one individual to the next, DNA

identification devices use the unique DNA to create a unique arrangement for every individual.

Cells in the body, whether collected from cheek cell, skin cell, blood cell, or other tissue share the same DNA. This DNA is exclusive for each individual (except for identical twins who share the same DNA pattern) and thus makes the identification process easywhen two samples are compared.

Just over thirty years ago, Sir Alec Jeffrey's positioned the foundation stone of modern molecular forensic sciences with the discovery of hypervariable minisatellites and DNA fingerprinting. Before that, accurate human individualization was not possible and the best that could be accomplished by forensic scientists was an exclusion probability grounded on data from gene product analysis of polymorphic blood groups and protein loci or RFLP. In contrast, with the exception of monozygotic twins, DNA fingerprints have the capacity to compare a sample to a unique individual and this capability to possibly individualize altered the mind set of forensic scientists forever. Advancement in forensic investigation has been prompt since the application of multilocus probes, through single-locus VNTR probes and microsatellite loci to arriving approaches including SNPs and proteomic microarrays.

3. Techniques

3.1 Restriction Fragment Length Polymorphism (RFLP)

Restriction Fragment Length Polymorphism was the first technique established to examine variable lengths of DNA fragments formed through DNA digestion. It uses the variations in DNA sequences due to the different locations of restriction enzyme sites. The method practicesthe use of restriction end nucleases to digest the DNA by cutting it at precise sequence arrangements. The resultant restriction fragments are then separated in gel electrophoresis and then shifted to a

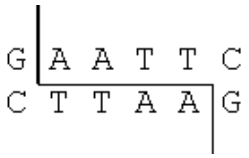
membrane using the Southern Blot technique. The separated DNA fragments are transferred and probe hybridisation is used to detect the fragments.

In 1980, Wyman and White discovered the first hyper variable locus present in human DNA. The allelic forms identified at this locus fluctuated in size and were so-called restriction fragment length polymorphisms (RFLP's). This procedure became the substance for more exhilarating discoveries.

In 1985, Dr Alec Jeffery's, using this technology, obtained a DNA fragment that when used as a DNA probe against genomic DNA of a human, attached unambiguously to a number of diverse RFLP sites on the genome generating a distinctive and reproducible banding pattern that was particular for an individual (with the exception of identical twins). He realised this banding pattern was inherited, with half the bands donated from the mother and half from the father. This procedure was named DNA fingerprinting. Originally, DNA fingerprinting was used to define family relationships in immigration applications. However, in September 1986, Dr Jeffery's was approached by police to use the genetic fingerprinting technology to assist a rape/murder investigation. The highly publicised case, Narborough Murder Enquiry, was the first murder investigation resolved by DNA fingerprinting.

3.1.1 Restriction Enzymes

The technique of DNA fingerprinting needs the DNA to be cut up into small fragments. Restriction enzymes perform this digestion. Restriction enzymes were discovered in bacteria, which practice them as a defence mechanism to cut up the DNA of other bacteria or viruses. Hundreds of different restriction enzymes have been extracted. Each one of them cuts DNA at a specific base sequence viz. *Eco*RI always cuts DNA at GAATTC as indicated below.



Other restriction enzymes cut at different sites, some examples are listed below.

Enzyme	Cutting Site
Pst I	CTGCAG
Bam HI	GGATCC
Hinf I	GANTC
Hae III	GGCC

In RFLP analysis, restriction enzymes cut up the DNA of an organism into fragments. A large number of short fragments of DNA are produced.

Restriction enzymes cut at the same base sequence always. Since, no two individuals have identical or duplicate DNA; no two individuals would have the same length fragments. For example, the enzyme *Eco*RI always cuts DNA at the sequence GAATTC (see above). Different people have different numbers of this particular sequence and will consequently have different fragment lengths. Additionally, some of the restriction sites will be at different locations on the chromosome.

3.1.2 The experimental steps being used in forensics laboratory for DNA profile analysis are as follows:

3.1.2.1 DNA extraction

It is possible to extract DNA from almost any human tissue. Sources of DNA found at a crime sight might include tissue from a deceased victim, blood, cells in a hair follicle, semen and even saliva. The

extracted DNA from the items of evidence is compared to DNA extracted from reference samples i.e. known individuals, normally from blood.

3.1.2.2 Digestion of DNA with a restriction endonuclease

Extracted DNA is exposed to a restriction endonuclease (Fig. 2), which will cut double stranded DNA whenever a specific DNA sequence occurs. The enzyme HaeIIIis most commonly used for forensic DNA analysis, which cuts DNA at the sequence 5'-GGCC-3'.

3.1.2.3 Agarose gel electrophoresis

The resulting DNA fragments are separated based on size factor via electrophoresis in agarose gels followed by DNA digestion.The migration rate of DNA fragments is slowed by the matrix of the agarose gel. The smaller DNA fragments move faster through the pores of the gel matrix than larger DNA fragments. As a result, a continuous separation of the DNA fragments according to size is obtained with the smallest DNA fragments travelling the greatest distance away from the origin.

3.1.2.4 Preparation of a Southern blot

After the completion of electrophoresis, the separated DNAs are denatured while still in the agarose gel by soaking the gel in a basic solution. Following neutralization of the basic solution, the single strand DNA molecules are transferred to the surface of a nylon membrane by blotting. After the inventor, Edwin Southern, this denaturation/blotting procedure is known as a Southern blot.The blotting of DNA to a nylon membrane conserves the spatial arrangement of the DNA fragments that occurred after electrophoresis.

3.1.2.5 Hybridization with radioactive probe

A single locus probe is a DNA or RNA sequence that is able to hybridize (i.e. form a DNA-DNA or DNA-RNA duplex) with DNA from a specific restriction fragment on the Southern blot. The duplex formation depends on the complementary base pairing between the DNA on the Southern blot and the probe sequence. The single locus probes are

usually tagged with a radioactive label for an easy detection, and are chosen to detect one polymorphic genetic locus on a single chromosome. Under conditions of temperature and salt concentration that favour hybridization, the Southern blot from step 4 is incubated in a solution containing a radioactive, single locus probe. The unbound probe is washed away after hybridization; consequently, the only radioactivity remaining bound to the nylon membrane is associated with the DNA of the targeted locus.

3.1.2.6 Detection of RFLPs via autoradiography

Autoradiography detects the locations of radioactive probe hybridization on the Southern blot. In this technique, the washed nylon membrane is overlaid by a sheet of X-ray film in a light tight container. The X-ray film records the locations of radioactive decay. After exposure and development of the X-ray film, the resulting record of the Southern hybridization is termed an autoradiograph, or authored.

3.1.2.7 Re-probe southern blot with additional probes

After an authored has been developed for the first single locus probe, the radioactivity on the Southern blot can be washed away with a high temperature solution, leaving the DNA in place. The Southern blot can be hybridized with a second radioactive single locus probe, and by repetition of steps 5-7, a series of different single locus probes. The set of autorads from a Southern blot is known as a DNA profile.

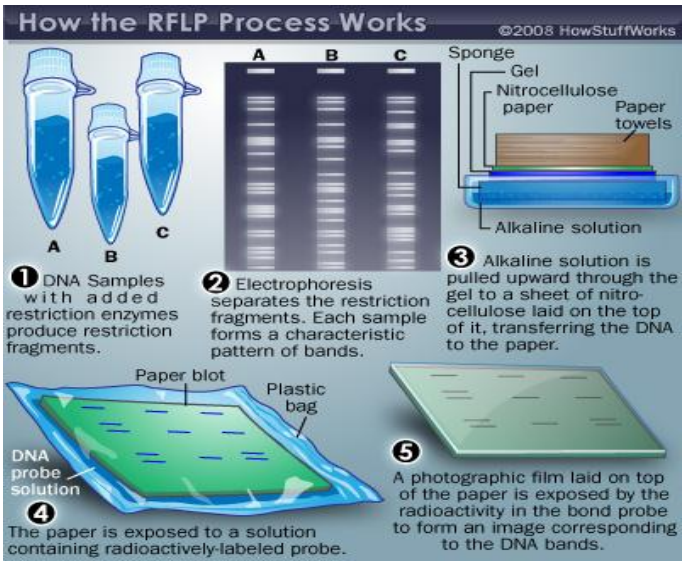

How the RFLP Process Works
©2008 HowStuffWorks
A
B
C
Sponge
Gel
Nitrocellulose paper
Paper towels
Alkaline solution
1 DNA Samples with added restriction enzymes produce restriction fragments.
2 Electrophoresis separates the restriction fragments. Each sample forms a characteristic pattern of bands.
3 Alkaline solution is pulled upward through the gel to a sheet of nitro-cellulose laid on the top of it, transferring the DNA to the paper.
Paper blot
Plastic bag
DNA probe solution
4 The paper is exposed to a solution containing radioactively-labeled probe.
5 A photographic film laid on top of the paper is exposed by the radioactivity in the bond probe to form an image corresponding to the DNA bands.

Figure2: Outline for procedure of Restriction Fragment Length Polymorphism.

3.1.3 Limitations:

However, DNA analysis with RFLP required comparatively large amounts of DNA and degraded samples could not be examined with precision. More effective, quicker and economical DNA profiling techniques have been developed, so RFLP is mostly no longer used in forensic science.

3.2 Amplified Fragment Length Polymorphism (AMP-FLP)

A range of DNA extraction methods has been used for forensic DNA analysis. For example, digestion of body fluid stains using SDS and proteinase K, followed by purification of DNA by extraction with phenol/chloroform and ethanol precipitation, is very successful and is routinely used for forensic samples analyzed by RFLP typing. However, this method had limitations when applied to a PCR-based DNA typing

method for forensic analysis viz. HLA DQα typing. For bloodstains, it was observed that, although adequate DNA was obtained for analysis, it could not always amplify using PCR. This failure in amplification process is found to be caused by the existence of hematin in bloodstains, since hematin is an inhibitor of PCR.

Another PCR-based DNA typing method, used for the analysis of amplified fragment length polymorphisms (AMP-FLPs) could be implemented in forensic laboratory but it was advantageous to assess a number of DNA extraction methods to decide the most suitable one for AMP-FLP analysis. The AFLP method was developed in 1995 by Vos *et al*. and has been used for numerous years in research laboratories and for patent applications.Factors considered, when various methods were compared include the yield of DNA, the suitability of DNA for amplification, existence of fragments of DNA on a silver stained gel and the differential amplification of alleles having different sizes in a sample. A variety of extraction methods was experimented for all these factors, including Chelex100 extraction, organic extraction followed by either ethanol precipitation or Centric on 100® dialysis and concentration and several commercially available DNA extraction kits.

The features of optimized AFLP analysis project the assay valuable for a number of clinical applications. The human identity testing has evolved from agarose gel-based separation of DNA restriction fragments to capillary electrophoresis platforms usage and this move has greatly improved the resolution of the separation technique.

AFLP is an excellent technology to be used in the detection, separation, and ascriptionof a microbial strain in the case of a bio crime. Several forensic cases include plant evidence that may be valuable to link a victim, a suspect, a vehicle, a weapon and crime scenes. With the introduction of novel DNA technologies, plant DNA material can be chemically extracted and typed using a multilocus detection method

called AFLP. It is a PCR-based method to produce DNA fingerprints and speedily screen genetic diversity. AFLP uses a pair of restriction enzymes to cut up the genomic DNA unlike markers such as microsatellites, where designed primers target the markers in the genome. Then, a pair of synthetic DNA fragments, adaptors, is attached to the complementary sticky ends of genomic DNA fragments during the ligation phase. It is followed by pre-selective and selective PCR stages, which use primers that match the known adaptor sequence along with additional selective nucleotides, which increase the specificity of amplification, thus reducing the total number of final fragments or loci. In the selective PCR stage, one of the primers having a fluorescent label attached allows the DNA fingerprints to be visualized by electrophoresis using a sequencer. Fig.3 shows the use of the adaptors/primers Eco and Pst and loci derived from 6 primer combinations i.e. selective primers with different selective nucleotides.

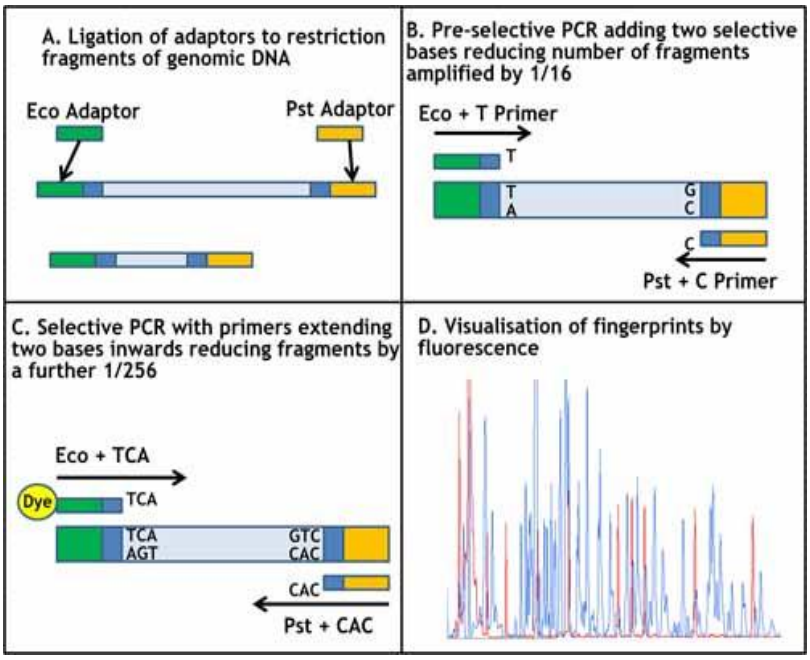


Figure3: Simplified diagram of the AFLP method following digestion of genomic DNA using Eco and Pst restriction enzymes. Base combinations are purely for illustrative purposes.

3.3. Polymerase Chain Reaction (PCR):

During the investigation of a crime, the amount of DNA evidence procured is often very small, thus for efficient DNA profiling, amplification is ideal. PCR allows for the exponential amplification of DNA fragments to the lengths of approximately 10,000bp. PCR is particularly helpful in the amplification of small amounts or degraded samples. In the 1980s, when RFLP was being developed for forensic use, PCR was developed for forensic applications. In 1983, a biochemist, Dr Kary Mullis adopted PCR to amplify DNA fragments of forensic interest in an automated process. He was awarded the Nobel Prize for his discovery

in 1993. PCR solved one of the major struggles facing by forensic biologists. RFLP typing required relatively large quantities of DNA.

3.3.1 The PCR Process

PCR methods are very susceptible to contamination by foreign DNA. On this ground, DNA extractions are always performed in a location physically isolated from the desk/lab where the subsequent amplifications would be performed.

A PCR reaction needs a number of fundamental primary components. Oligonucleotide primers complementary to the DNA target to mark the target to be amplified, with two primers in use. Fluorescent tags are sometimes added to the primers to locate amplified DNA in electrophoresis. DNA polymerase enzyme adds nucleotides to the 3' end of the primers and allows the DNA strand to be copied. Other components required include a template DNA, a reaction buffer with $MgCl_2$ to ascertain ideal conditions for the performance of the DNA polymerase enzyme and deoxyribonucleotides to build the DNA molecule.

The three steps of PCR process take place within a thermal cycler, capable of reaching and upholding preset temperatures very accurately. The DNA samples are supplemented to a reaction buffer, a salt solution that is buffered at the optimal pH so that the polymerase enzyme can function properly. The building blocks of DNA, four nucleotides, are added to the buffer along with the *Taq* polymerase enzyme that helps in catalysing the extension step.

3.3.2 The steps in the PCR reaction are as follows:

3.3.2.1 Denaturation

After adding the DNA to the PCR tube containing the reaction mixture it is heated to 95°C which leads to the denaturation of the double-stranded DNA. The H-bonds present between the base pairs

break and single-stranded DNA is formed. Each strand becomes the template for the formation of a new double-stranded DNA.

3.3.2.2 Annealing

In this step, a short strand of synthetic DNA, primer, attaches to each of the separated strands. They mark the starting points for the addition of new bases to start the replication of each strand. The temperature of thermal cycler drops to 60°C for this step.

3.2.2.3 Extension

Under the influence of Taq polymerase at the temperature of 72°C, single bases/ nucleotides are added to the primer, one by one. At each of the complementary single strands created by the denaturation process, this process occurs so at the end two identical pieces of double-stranded DNA are produced. This makes one PCR cycle. The temperature is raised once again to 94°C and the process repeats in this way .Four single strands are formed after the denaturation of the two strands. They again go through annealing and extension, which forms four new double strands. The process continues till 25–40 cycles, which takes nearly 2-3 hours. This produces approximately one billion copies of the original DNA, sufficient for additional typing. The steps in PCR are shown in Fig.4.

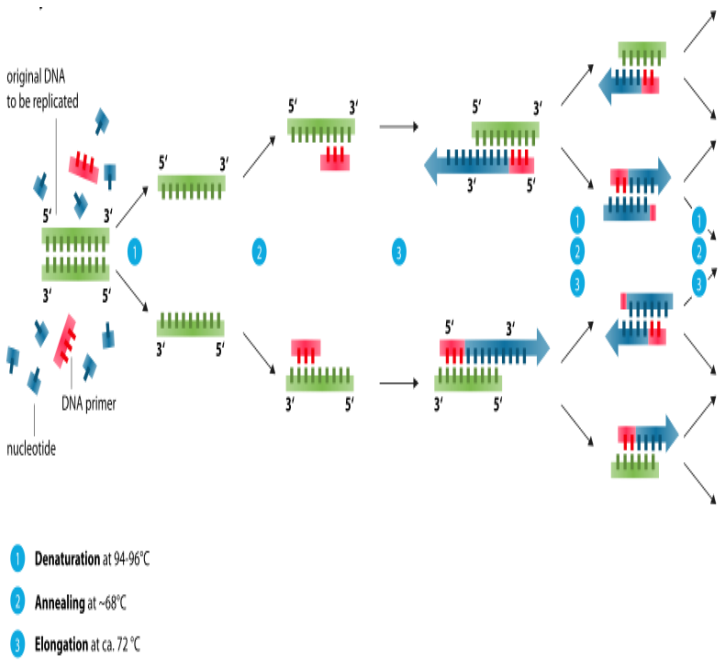


Figure 4: Exponential amplification by Polymerase Chain Reaction

3.3.3 Basic PCR Protocol

The protocol serves as a wide-ranging guideline and a preliminary point for any PCR amplification. Optimal reaction conditions (primers, $MgCl_2$, incubation times and temperatures, concentration of Taq DNA Polymerase and template DNA) may differ and need to be calculatedfor the experiment.

1. Add the following components to a DNase/RNase-free 0.5-ml microcentrifuge tube placed once. Scale the reaction volumes as needed. Prepare a master mix for multiple reactions, to minimize reagent loss and to enable accurate pipetting.

Components	Volume	Final Concentration
10X PCR buffer minus Mg	10 µl	1X
10 mMdNTP mixture	2 µl	0.2 mM each
50 mM MgCl2	3 µl	1.5 Mm
Primer mix (10 µM each)	5 µl	0.5 µM each
Template DNA	1-20 µl	-----
Taq DNA Polymerase (5 U/µl)	0.5 µl	2.5 units
Autoclaved distilled water to 100 µl		

2. Mix contents of tube and overlay with 50 µl of mineral or silicone oil.
3. Cap tubes and centrifuge briefly to collect the contents to the bottom.

Before the amplification of DNA in a thermal cycler, the DNA must be extracted from the collected sample from crime scene. Several steps need to be taken to isolate DNA from the cells, the cell membrane and nucleus are broken to expose the DNA without destroying it. This can be achieved through numerous different extraction methods. Chelex, created by the Bio-Rad company, is a resin that is added to a sample of DNA. Chelex resins being negatively charged help to remove positive metal ions. Chelex resins bind Mg+ ions in order to prevent DNA nucleases from becoming activated. After the centrifugation of the sample, purified DNA can be removed from the supernatant.

Other methods of extracting DNA include specialized cellulose paper called FTA™, the use of organic chemicals (phenol-chloroform) and differential extraction. All achieve the same purpose.

3.3.4 DNA Typing of PCR Product

The products aredetected after the amplification process is complete. Since the product is more relative to that produced by RFLP, it is not compulsory to employ very delicate detection methods such as radioactive labelling. A yield gel experiment can be performed on agarose and stain the product with EtBr.

3.3.4.1 Variable Number of Tandem Repeats (VNTRs)

The DNA sequences that create the variability contain runs of short, repeated sequences, such as GTGTGT..., which are presentin various positions (loci) in the human genome. The number of repeats in each run is extremely variable in any population, ranging from 4 to 40 in different individuals. This type of segment of repeated nucleotides is generally referred to as a hypervariablemicrosatellite sequence, also known as a VNTR (variable number of tandem repeats) sequence (Fig.5). Individuals usually inherit a different variant of each VNTR locus from their mother and from their father because of the variability in these sequences; two unrelated individuals therefore do not generally contain the same pair of sequences. A PCR reaction using primers produces a pair of bands of amplified DNA from each individual, one band representing the paternal variant and the other representing the maternal variant(Fig.6A). The length of the amplified DNA and its position after electrophoresis depend on the exact number of repeats at the locus.

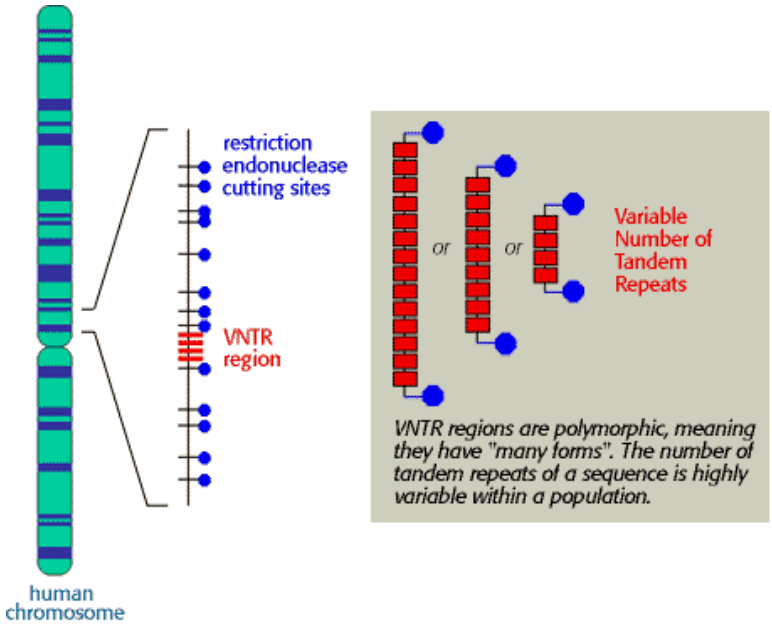


Figure 5: Variable number of Tandem Repeats with RE cutting sites

In the schematic example shown in Fig.6B, the three VNTR loci are analyzed from three suspects (individuals A, B, and C), producing six bands for each person A,B and C after polyacrylamide gel electrophoresis. Nevertheless, some individuals have numerous bands in common; the overall pattern is quite unique for each. The band pattern can therefore serve as a fingerprint to identify an individual nearly uniquely. The fourth lane, F, contains the products of the same PCR reactions carried out on the forensic sample. The starting material for such a PCR reaction can be a single hair, scratched tissue or a tiny sample of blood that was left at the crime scene. The more loci that are examined, the more confident one can be about the results. While examining the variability at 5-10 different VNTR loci, the odds that two random individuals would share the same fingerprint by chance are

approximately one in 10 billion. In the case shown here, individuals A and C can be eliminated from inquiries, while B remains a clear suspect. A similar approach is now used routinely for paternity testing.

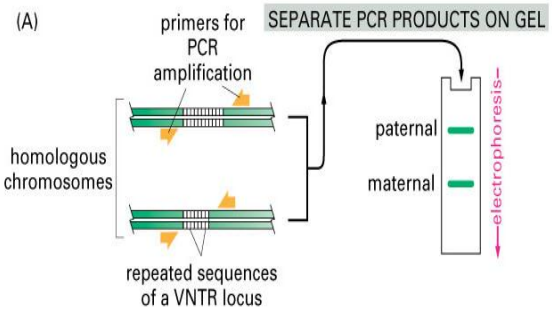


Figure 10-30 part 1 of 2 Essential Cell Biology, 2/e. (© 2004 Garland Science)

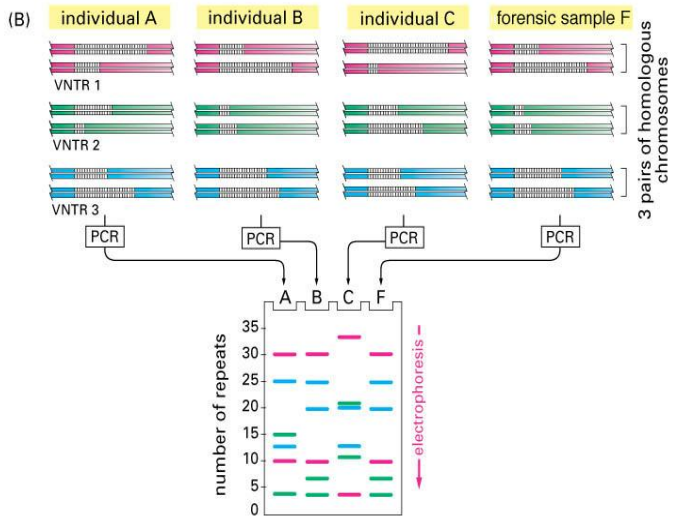


Figure 10-30 part 2 of 2 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Figure 6 (A, B): Analysis of VNTR loci from three suspects

3.3.5 The application of PCR technology to forensic science

For the PCR to be a useful application in forensic science the selected target sequences should carry several allelic variations, possibly all less than 1000 bp in size. A range of polymorphic loci that fits for biological discrimination purposes fulfils these parameters. The best known of the PCR systems used for forensic science are the D1S80 and the HLA DQA1systems.

The HLA DQA1 kit uses a series of allele specific probes to identify the various allelic forms.

D1S80 is a VNTR locus where the allelic forms are simply separated and identified by size. The term AMP-FLP's has been used to define D1S80 and other VNTR loci that can be amplified by PCR. Other AMP-FLP's that have also proved to be useful PCR systems are the 3' HVR (Hyper variable repeats) region of the apolipoprotein B gene, D17S30 and Col1A2.

A third PCR system, the mitochondrial D loop region, has more recently fascinated a deal of attention by forensic scientists. This is a highly polymorphic DNA locus present on the mitochondrial genome. When the D loop locus is once amplified, it is best studied by direct sequencing of the PCR products. This system is likely to receive a great deal ofinterest with the aid of modern automated DNA sequences.

3.3.6 Advantages of PCR technology

1. It is more sensitive than the single-locus and multi-locus probe technologies and the protein and antigenic systems.
2. The PCR process can be tailored for a particular locus.
3. The technology is potentially very cheap.
4. The technology is simple to understand and much easier to perform than RFLP technology.
5. Radioactive isotopes are not required.
6. Results can often be obtained from crude DNA preparations.

7. The PCR products do not alter by degradation of the template DNA caused by decomposition of the sample.
8. The amplification process is largely automated, and therefore precise.

3.3.7 Disadvantages of PCR technology

1. False positives may occur if care is not taken to prevent the contamination of samples.
2. The PCR process may be inhibited by various chemicals present in the DNA extract viz.Haemoglobin, metabolites released by vaginal flora and various organic acids in soils are powerful inhibitors of the PCR process.
3. PCR errors like inability of proof reading of Taq DNA polymerase may result in error in amplification. Mispriming and excessive primer dimer formation may also occur.

4. Short Tandem Repeats (STRs)

STRs, sometimes referred to as microsatellites or simple sequence repeats (SSRs), are found as short sequences of DNA, length of 2-5 base pairs, repeated many times in a head and tail manner, viz. the 20bp sequence of "GATAGATAGATAGATAGATA" would represent 5 head and tail copies of the tetramer "GATA". The dissimilar number of copies of the repeat element in a population leads to the polymorphisms in STRs.DNA fingerprinting relies upon the analysis of these short tandem repeats (STRs).Only few STR markers, which express a high degree of polymorphism, making them of specific use, are used in forensic DNA profiling.

STR markers are generally of three types – simple (identical length repeats), compound (two or more neighbouring repeats) or complex (numerous different length repeats). They are found on autosomal and allosomal chromosomes. Inaccuracy of DNA polymerase

enzyme in copying the region causes the variability in STRs.Tetra- and pentad-nucleotide repeats are preferred for STRs used in forensic science as they provide a high degree of fault free data and suffer less environmental degradation

Commercial kits are available to produce DNA profiles containing these main STR loci (Table 1).

Table 1. Characteristics of the 15 STR Loci Present in the Commercially Available Kit AmpF/STR Identifiler

STR Loci	Chromosomal Location	Repeat Motif	Allele Range	PCR Product Sizes inIdentifiler Kit (dye label)
CSF1PO	5q33.1	TAGA	6–15	305–342 bp (6-FAM)
FGA	4q31.3	CTTT	17–51.2	215–355 bp (PET)
TH01	11p15.5	TCAT	4–13.3	163–202 bp (VIC)
TPOX	2p25.3	GAAT	6–13	222–250 bp (NED)
VWA	12p13.31	[TCTG] [TCTA]	11–24	155–207 bp (NED)
D3S1358	3p21.31	[TCTG] [TCTA]	12–19	112–140 bp (VIC)
D5S818	5q23.2	AGAT	7–16	134–172 bp (PET)
D7S820	7q21.11	GATA	6–15	255–291 bp (6-FAM)
D8S1179	8q24.13	[TCTA] [TCTG]	8–19	123–170 bp (6-FAM)

D13S317	13q31.1	TATC	8–15	217–245 bp (VIC)
D16S539	16q24.1	GATA	5–15	252–292 bp (VIC)
D18S51	18q21.33	AGAA	7–27	262–345 bp (NED)
D21S11	21q21.1	[TCTA] [TCTG]	24–38	185–239 bp (6-FAM)
D2S1338	2q35	[TGCC] [TTCC]	15–28	307–359 bp (VIC)
D19S433	19q12	AAGG	9–17.2	102–135 bp (NED)
Amelogenin (sex-typing)	Xp22.22 Yp11.2	Not applicable	Not applicable	X = 107 bp (PET) Y = 113 bp (PET)

4.1 Performing STR Analysis

This method differs from RFLP since in STR analysis DNA is not cut with restriction enzymes. Probes are attached to preferred regions on the DNA, and a PCR is employed to discover the lengths of the short tandem repeats.

The whole process for STR typing comprisesof collection of sample, extraction of DNA, quantisation of DNA, amplification of multiple STR loci by PCR, separation and sizing of STR allele, STR typing followed by profile interpretation, and possibly a report of the statistical significance of a match. Current forensic systems apply 10 (e.g. United Kingdom) or 13 (e.g. United States) STR loci. Kits with PCR primers for the standard STR loci are available commercially.

Numerous PCR reactions are performedconcurrently in a single tube at different STR loci, which give several products (two for each locus). The required components are as follows:

- A DNA sample, e.g. blood or buccal cells from a suspect or a tissue, hair, nail from the scene of a crime
- Two oligonucleotide PCR primers: one unlabelled reverse primer and one primer labelled at the 5′-end with ^{32}P
- A DNA polymerase (thermos table)
- Four deoxynucleoside triphosphates which are as follows: dATP, dGTP, dCTP, dTTP.

The labelled PCR products separate according to size when run on a polyacrylamide gel. DNA ladder is obtained and works as characteristic of an individual (Fig.7).

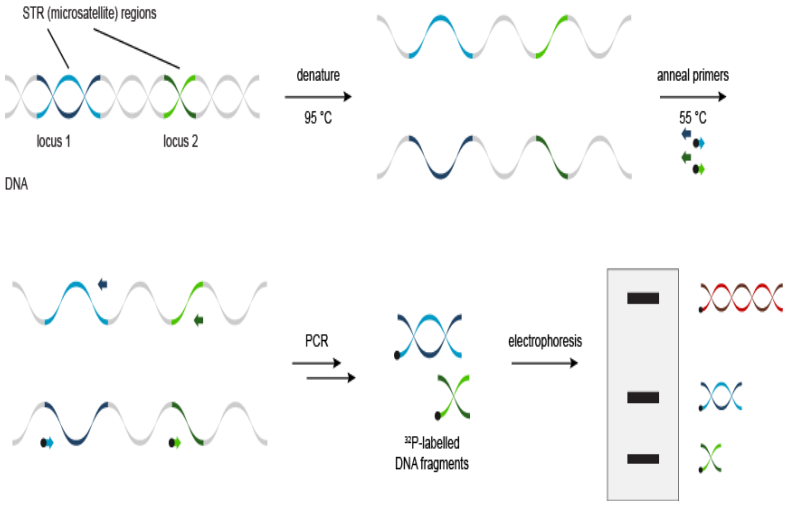


Figure 7: DNA Fingerprinting by STR analysis

4.1.2 Fluorescent STR analysis

In an advanced variant of STR analysis, PCR primers are labelled with different fluorescent dyes(Fig.8).Still, only a limited number of fluorescent dyes with adequate spectral characteristics have been developed, three different types of fluorescent dyes are used.The use of different fluorescent labels facilitates the distinction between the products and bands originating from different STR loci.

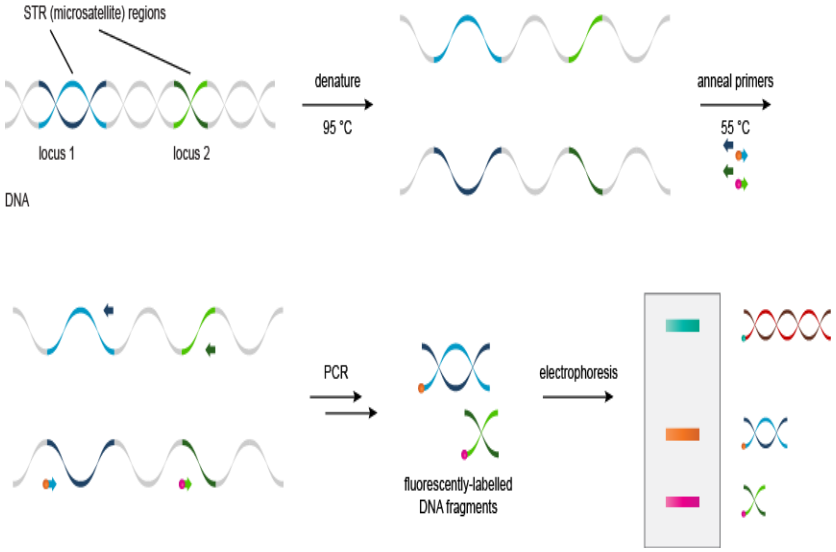


Figure 8: Fluorescent STR analysis

4.2 Paternity testing

DNA paternity testing helps in parental dispute cases. DNA paternity test is either an inclusion (the alleged father is regarded as the biological father with 99.9% accuracy) or an exclusion (the alleged father is not the biological father with the accuracyof 100%). The tested parties include the alleged father, the mother and a child.

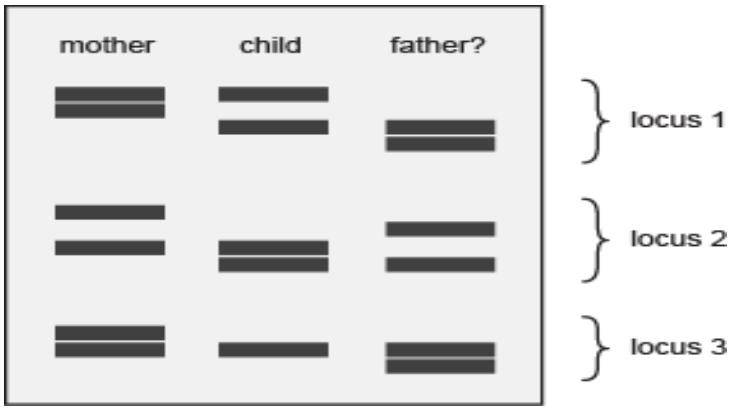


Figure 9: STR analysis in a parentage dispute. An example of STR analysis in a parentage dispute (using 3 STR loci).

4.3 Benefits of STR Analysis

Presently, it is a standard method for genetic profiling. STR analysis used today includes four to five nucleotide repeats. It allows scientists to achievecomparativelyaccurate information.

In the September 11, 2001 attack on the World Trade Centre (WTC), New York, thousands of people were killed.Forensic scientists faced the difficulty of ascertaining victims with degraded samples and required advanced and refined forms of DNA analysis.

4.4 Value of Technology to Identify Victims

In this comparatively newer approach, the concept involved smaller STR assays. The advantage of the smaller size is that a damaged DNA sample will still have fragments in the thirteen locations that differ among humans that are usable and can be analysed.

4.4.1 Understanding the DNA Analysis Used for September 11 Victims

These small fragments are called as DNA primers which bind to particular sequences of base pairs. The primers helps scientists procure

some of the DNA fragments to copy it to provide a sample of anadequately large size.

5. **Y-Short Tandem Repeats**

Y- STRs are present on the Y chromosome (Fig.10). Generally found on the short arm of the Y chromosome, the coding genesare important for sex determination of males and spermatogenesis. These Y-STRs are polymorphic among unrelated males.They are inherited through the paternal line with minute change through several generations.

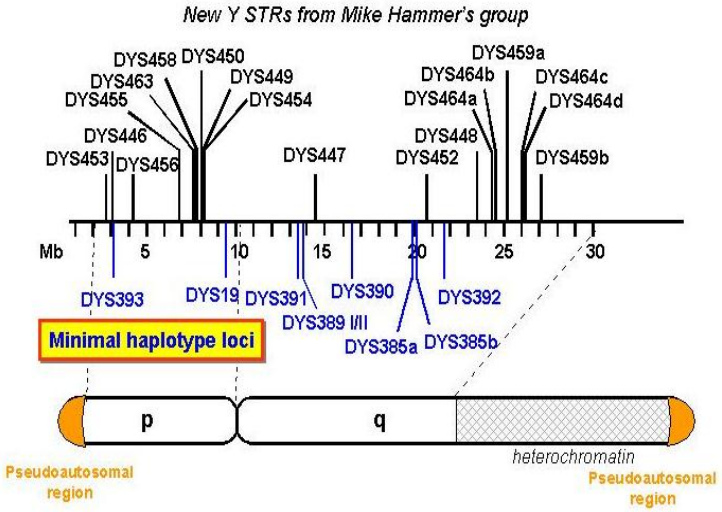


Figure 10: Y STR Positions along Y-Chromosome

5.1 Y-STR Testing

It focuses exclusively on the male DNA present in a particular sample. This technique can completely ignore the presence of female

DNA because this technique amplifies definite DNA segments that are positioned only on the Y chromosome.

5.1.1 Application of Y-STR testing

When the forensic results come back specifying that little male DNA present in the sample collected from crime scene is masked by the abundance of the victim's (female) own DNA then it could be further analyzed with Y-STR typing. In other cases viz. rape cases where sperm quantity is absent due to either a vasectomised rapist or a long duration between the alleged violence and gathering of the rape kit samples or bite mark swabbing procured from a female victim can produce Y-STR profiles even if large amounts of DNA from the female victim herself is present. In gang rapes, where several males may have contributed to the collected samples-STR typing can be helpful. This testing can to knowthe exact no. of males involved in the crime.

5.1.2 Disadvantages to Y-STR testing

There are few disadvantages to this type of testing which are as follows:

Since, all males of a family would have exactly the same Y-STR profile any possible male suspect's Y-STR profile will be matching his father, his brothers, his uncles' Y-STR profile. Additionally, it can be possible for individuals to have the same common Y-STR profile without being closely related. The advanced available Y-STR kits in use today concentrate upon either 12 or 17 loci on the Y chromosome to solve this issue. If the profiles of different suspects do not match, it is 100% exclusion and then other potential suspects can be focused. If the profiles match, this can be understand that the investigation is in correct direction.

Another possible disadvantage is that the results of Y-STR typing cannot be uploaded into the National DNA Database (CODIS) presently.

5.2 Additional forensic uses of Y-STR testing

There are plentiful other forensic uses for Y-STR analysis, comprising paternity cases and missing persons cases.

5.2.1. Case 1: paternity testing

A woman, her child and the accused possible father of her child were profiled using 15 autosomal STR markers. In the child and alleged father, the Y STR profiles were identical. Thus, the result concluded the alleged father being the biological father.

5.2.2. Case 2: sibling confirmation

A mother and her two sons were referred to the lab to assess sibling ship. The father was unavailable for testing.Y STR testing with markers demonstrated that the putative sons carry a Y chromosome sharing an identical Y haplotype at all 11 markers used.

6. X-chromosomal short tandem repeats (ChrX STRs)

X-chromosome short tandem repeat (X-STR) markers (Fig.11) have lead over autosomal and Y chromosome markers in kinship cases. It is used where the alleged father is absent and the child is female.Female individuals receive one X from the mother and the other X from the father while male individuals inherit one X-Chr from mother. Thus, female individuals from same father share their paternal chromosome X.

X-STRs are particularly helpful in paternity testing and kinship analyses, such as mother-son, father-daughter, kinship testing of putative sisters and grandmother-granddaughter kinship testing. When second and third degree kinships are considered, extremely polymorphic STRs are required.

A girl who disappeared for several years, X-STRs were used for the identification.A man suspected to be the murderer of another woman, was found to have a head scarf similar to the one belonging to the girl in his house. Inside the head scarf some hair was found.The

relationship between the hair found and the mother and the sister of the girl was verified in the absence of any biological sample.

For the identification of a biological material supposed to be belonging to a girl who disappeared for several years X-STRs were used.In fact in the house of a man suspected to be the cause of another woman's murder, a head scarf similar to the one belonging to the girl and inside it some hair was found. In the absence of any biological sample belonging to the girl who disappeared, the relationship between the hair above and the mother and the sister of the girl was verified. DNA typing of hair exposed that all of them were from a female and that they disclosed the same X-STR profile.The present case establishes that X-STR markers are beneficial even in special reverse paternity test.

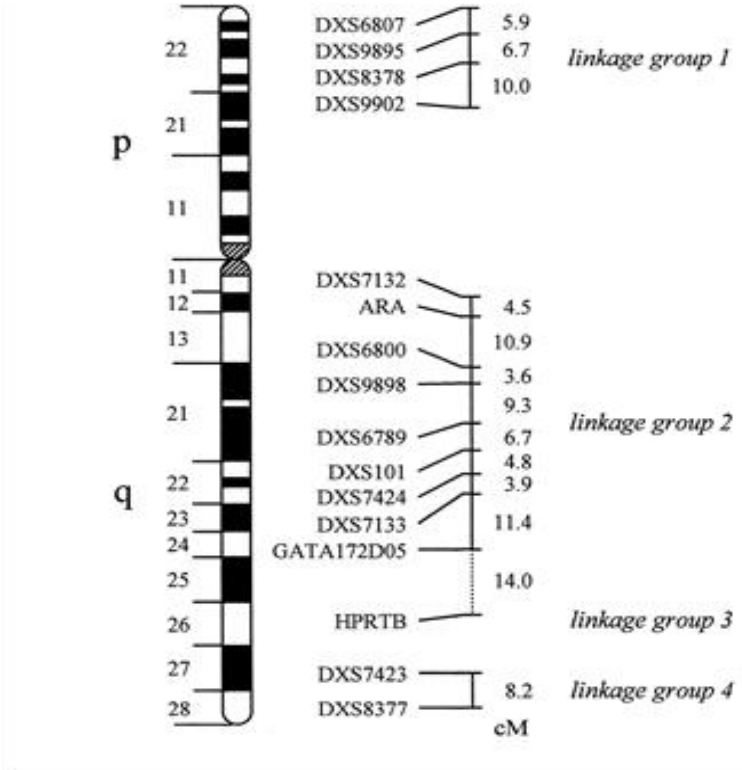


Figure11: Localisation of ChrX STRs used in forensic practise.

The order and approximate position of STRs on the ChrX ideogram is based upon publicised map data (Marshfield, NCBI) and RH and genetic mapping studies. Pair-wise genetic distances (in cM) are calculated from maximum likelihood estimates of pair- wise recombination fractions using the Kosambi mapping function.

6.1 Power of ChrX markers in stain analysis

With a few exemptions, AS (Autosomal) markers are more powerful than ChrX markers. In a mixed stain of male and female, the probability of all male alleles being incorporated in the female component is greater for ChrX than for AS markers. So, it is not suitable

to use ChrX markers to assess male traces where female contamination is present. But, to identify female traces present in male contamination, ChrX markers are more competent than AS markers since the female alleles can be fully included in the male component only if the female is coincidentally homozygous at all loci. Examination of female traces on a male rape suspect with the help of ChrX typing has already been performed efficaciously in practice.

7. mtDNA in forensic science

The examination of mitochondrial DNA is important for the identification of skeletal remains like teeth and bones. The average number of mtDNA molecules is 5000 in the epithelial cells, preferentially used in forensic case work. Due to this high copy number, it helps in better sensitivity for detection in cases where extremely small quantity of DNA or extremely degraded DNA is obtainable. Since, sperm mtDNA is degraded just after fertilization; the complete set of mitochondrial DNA in all the cells of all individuals is derived from mother. Due to this maternal inheritance, maternal family relationships can be determined bythe analysis of mtDNA even when a gap of several generations exists between a living person and an ancestor.

mtDNA testing may effectively progress the investigation and prosecution of cases with inadequate biological evidence, such as degraded skeletal remains and telogen hairs. mtDNA testing is also helpful in the area of post-conviction relief.

The most commonly used technique for regular forensic analysis of mtDNA is Sanger sequencing of two hypervariable regions (HVI and HVII) (Fig.12) within the control region (Displacement Loop, “D-loop”). It takes weeks to obtain reliable results. Other methods include minisequencing and the use of sequence-specific oligonucleotide probes and Pyrosequencing™ technology.

7.1 Forensic Applications

7.1.1 Mitochondrial DNA analysis is an appropriate method for:

- Charred remains
- Degraded specimens
- Old skeletal and fingernail samples
- Hair shafts

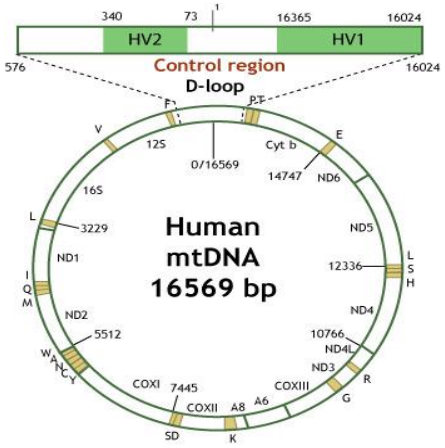


Figure12: Human mitochondrial DNA showing hypervariable regions

mtDNA is useful for the remains recovered from a missing person or in cases of mass disaster. The remains are found often in highly fragmented stage or only miniscule sample sizes are collected, such as a sliver of a bone or a single tooth. Biological material of known maternal relatives who are even quite distant could be used as reference sample for direct evaluation for the recovered remains. Next Generation Sequencing (NGS) can help in forensic testing to meet the challenges faced by mtDNA practitioners. Since, mixture interpretation is possible with this technology; NGS could broadenthe array of samples suitable for mtDNA analysis.

7.1.2 Other Applications

Scientists have used mtDNA analysis for medical studies, evolutionary studies, migration studies, genealogical studies, and historical identifications.

8. Microbial Forensics

The anthrax bioterrorism scare in 2001 that followed 9/11 Twin Towers attack further decreased America's perception of safety. Consequently, the focus on bioterrorism has increased and has become a great concern.

Biological warfare is an ancient experience. Some examples in recorded history take us back to the 6^{th} century BC. The Assyrians seemingly poisoned the wells of their opponents with rye ergot. In 1754-1763, during the French and Indian war, the English captain, Ecuyer, offered blankets smeared with smallpox to Indian loyal to the French.The exercise of biological weapons became more ubiquitous in the 20th century. During World War I the Germans tried to spread cholera and plague in Italy and St. Petersburg.

8.1 Molecular techniques to identify pathogenic organisms

There are several methods to recognize pathogenic microorganisms. These techniques are based on analyses as diverse as antigens, peptides and other ligands for identification. The key disadvantage of these methods is that they require the distinctness required to offer a positive detection. The genetic composition of all the organisms is unique. Therefore, most of the methods today stresses on the use of genetic markers for recognition and identification of bio pathogens. Any polymorphic region found in an organism's genome which can help in positive identification is a genetic marker.

Several researchers have taken help of techniques which are based on length polymorphisms (e.g., AFLP) to effectively differentiate

different bacteria species but these techniques lack the preciseness required to differentiate microorganisms which are very closely related to each other.

Q-PCR has also been used to detect and identify different pathogenic microbial and fungal species. But, the main disadvantage of PCR is its limited multiplexing capability. The method does not compete the other technologies viz. DNA microarrays or bead-based assays since they have high multiplexing ability.

Microarrays are extensively used because of their specificity, sensitivity and multiplexing capabilities. A single array chip is generally imprinted with thousands of oligonucleotides. DNA chips are used to identify and differentiate several pathogenic bacteria in the *Proteus, Vibrio, Escherichia, Shigella, Salmonella, Mycobacterium* and *Bacillus* genera.Since, microarrays require experts in bioinformatics for their design as well as analysis; it becomes very expensive for routine use. There are other problems related with the immobilization of the probes on the chip. When compared to DNA Microarrays, liquid phase assays viz. the bead-based technologies are not as vulnerable to the thermodynamic issues allied with the probe-target binding.

Liquid array analyses provide high sensitivity and high specificity, quantitative and multiplexing abilities.One of the unique bead-based liquid array systems comprises a convergence of flow cytometry and microsphere technology. In the microarray, fluorescent signals were often too feeble to analyse and had more cross-hybridization than those in the bead-based liquid array method.

Another powerful method in pathogen detection is ELISA. The difference between the Luminex assays and the ELISA is that ELISA is singleplexed and Luminexis multiplexed and more sensitive.

9. Non-human DNA Forensics

9.1 Introduction

Non-human DNA analysis in forensic science has gained very rapid growth in recent years. It has applications range from investigations of rape and murder of humans to cruelty and poaching in animal/wildlife species. DNA evidence from bacteria, animals, plants and viruses has been used in criminal investigations. When there is an absence of biological evidence to directly relate a suspect's DNA to a victim or crime scene, the facility to employ non-human DNA found as trace evidence at a crime scene to indirectly make this connection is a highly useful one. With human DNA profiling, individualization is most commonly endeavoured using STR genotyping. In some cases where STR profiling is challenging due to the lack of sufficient amounts of nuclear DNA, mitochondrial DNA (mtDNA), being more robust to degradation and available in larger copy numbers, has been effectively evaluated. In most of the animals, the 50 HVR region of the control region or D-loop is ordinarily sequenced, and nowadays a number of databases are present in domestic species for mtDNA haplotype frequencies. But like humans, there are limitations in using mtDNAas it is not suitable for positive individual identification. Additionally, its exclusion capacity is lower than that seen with STR loci. This problem is further deepened in domestic animals such as dogs and cats, which have far higher frequencies of common haplotypes and far fewer haplotypes than humans.

9.2 Types of non-human biological evidence

In addition to hair, other biological material found as trace evidence has been used efficaciously in forensic investigations e.g. DNA from saliva around bite wound or bitten material, DNA from faeces and urine of dogs and DNA from plant and soil evidence.

9.3 Types of investigations

9.3.1 Analysis of soil DNA

Identification of soil gathered as trace evidence can provide important hints leading to settlement of a case. Having an extremely complex matrix, soil characterisation is often carried out using DNA profiling of bacterial groups in the soil which uses the length hypervariability of 16s ribosomal RNA domains and the generated met genomic profiles are then compared between regions.

9.3.2 Analysis of viral DNA

In a criminal case State of Louisiana v. Dr. Richard Schmidt, 1998 in the USA, Dr Schmidt, a gastroenterologist, who had an extramarital affair with his nurse Ms Trahan for years, was suspected of her attempted murder. It was assumed that when their relationship was ended by Ms Trahan, Dr Schmidt injected Ms Trahan with blood from one of his HIV-positive patients. After having cleared other possibilities of infection, in order to ascertain that Ms Trahan had been infected by Dr Schmidt's patient, sequences of two HIV genes were acquired from both individuals and equated to other HIV patients in the local area. Phylogenetic analysis was performed. A close relationship between the HIV sequences of Dr Schmidt's patient and Ms Trahan was discovered, and those sequences from the local population sample were much more distant. This characterized first time that Dr Schmidt was convicted of attempted murder by using phylogenetic analyses in a criminal court case.

9.3.3 Drug enforcement

In drug prosecution cases, it is often problematic to recognize controlled substances, particularly when drugs are dried or powdered. Tsai *et al.* has effectively used DNA sequencing of the ribosomal internal transcribed spacer regions, ITS 1 and ITS 2, in the nuclear genome and the trnL-trnFintergenic spacer region in the chloroplast genome to

identify *Cannabis* concealed as lawfully imported plant products in Taiwan. Since *Cannabis* plants can be propagated either clonally by taking cuttings or from seed, DNA approaches competent to discriminate between the two can support investigations into finding sources of *Cannabis* growers and linking suspects to individual growing operations.

9.3.4 Analysis of animal STR DNA

The first known use of animal DNA evidence in a criminal investigation was in Canada in 1994. When the RCMP (Royal Canadian Mounted Police) found the dead body of Shirley Duguay who had been missing for 8 months from Prince Edward Island, they suspected that Douglas Beamish, her former husband, was involved in her murder. A blood stained leather jacket was found near Ms. Duguay's body, which was supposed to belong to Beamish. Although blood on the jacket was matched to match Ms. Duguay's, an identification of the jacket's ownership remained difficult.Forensic scientists found a small number of white hairs of a cat on the jacket. Beamish, living with his parents, owned a white cat called Snowball.Investigators took a blood sample from Snowball and referred it for DNA testing along with the hair from the jacket. The outcomes exposed that 10 dinucleotide STR loci had the similar alleles in both the DNA from the DNA recovered from the root of one hair found on the jacket and Snowball's blood. This evidence connected Beamish in Ms. Duguay's murder and he was later imprisoned.

9.3.5 Analysis of animal mtDNA

For dogs, a number of mtDNA control region haplotype frequency databases have been testified in various countries e.g. Japan, USA,Sweden, Belgium,UK with sequences ranging in length from 580 bp–1,4000 bp (complete) in length. Some haplotypes are found to be common in dog populations and other rare haplotypes are supposed to be important in terms of evidentiary value. State of California v. David

Westerfield, 2002 was first criminal case where inadequate DNA prevented STR analysis and compelled the use of mtDNA. When STR implication was ineffective, the mtDNA control region was used for implification and sequencing and a match was obtained between dog hair obtained from crime scene and Van Dam's dog hair. The frequency of the haplotype was assessed at 9% (Questgen Forensics). This was used as proof in court and David Westerfield was convicted of Danielle's kidnap followed by murder and sentenced to death.

9.3.6 Dog/bear attack

In some cases where humans are pounced on by dogs, DNA evidence are obtained from saliva around the bite wound and used to bring the owners of a brutal dog to justice.Frosch *et al.* reported in a rare case of a fatal attack of a human by a bear in Europe that the wrong bear had been shot by officials in Bulgaria. They used 12 STRs on bear hair collected from the scene of the attack and hair/tissue from the carcass of a shot bear to reveal this wrong action of management.

9.3.7 Plant DNA

The use of plants as evidence in criminal investigations is called Forensic botany.It comprises a number of spheres including, plant anatomy, plant systematic and palynology (pollen analysis). During the commission of crimes happened outdoor, plant material may be transferred from the crime scene to the perpetrators or victim. This plant material may help in be probation due to the individual genotype identified or restricted geographical distribution of the plant species. The major experimental tool of the forensic botanist remains the light microscope but the intra-species genetic variation is best decided by molecular genetics methods.

The very first reported application of plant DNA evidence reminds the molecular identification of seed pods from a Palo Verde tree used to relate a suspect to a precise crime scene. An Arizonan geneticist

designed multiple primer RAPD analysis and then was able to reveal inter-individual variation among diverse Palo Verde trees. The suspect was imprisoned.

9.3.8 Patent infringement

In a patent violation case, Congiu *et al.* confirmed using RAPDs that a patented variety of economically important strawberry (Marmolada) had been unlawfully commercialised by farmers in Italy. The RAPD results showed banding patterns in 13 of the 31 plants tested that were alike the Marmolada variety. This confirmation was used in court against the farmers.

9.3.9 Other

In the Cessna aircraft crash investigation in Oklahoma in 2008, Dove *et al*. used COI DNA sequencing on feathers procured from the engine in order to recognize the species of bird responsible for the crash. In a case of suspected suicidal poisoning of a 23-year-old female student, her body was found beside green vomit and red berries, an autopsy showed presence of partially crushed yew leaves in the stomach. Gausterer *et al.* confirmed presence of Taxus spp. in the stored gut contents using ribosomal ITS 1 DNA sequencing, providing evidence for suicidal poisoning.

9.4 Prospective

The prospective offered by the application of non-human DNA in solving crime cases is clearly gigantic. An enormous range of crimes has productively been solved using nonhuman DNA from varied species. Forensic laboratories normally necessitate accreditation for quality assurance purposes and this is an expensive process and unfeasible to acquire for many low-scale laboratories.

References:

1. Ballantyne, J. 1997. Mass disaster genetics. Nat. Genet. 15:329-331.
2. Budowle, B. and F.S. Baechtel. 1990. Modifications to improve the effectiveness of restriction fragment length polymorphism typing. Appl. Theore. Electrophor. 1:181-187.
3. Budowle, B. and R.C. Allen. 1987. Electrophoresis reliability. I. The contaminant issue. J. Forensic Sci. 32:1537-1550.
4. Budowle, B., B. Shea, S. Niezgoda, and R. Chakraborty. 2001. CODIS STR loci data from 41 sample populations. J. Forensic Sci. 46:453-489.
5. Bruce Budowle and Angela van Daal. 2008. Forensically relevant SNP classes. BioTechniques 44:603-610.
6. Budowle B and van Daal A. 2009.Extracting evidence from forensic DNA analysis: future molecular biology directions.Biotechniques 46: 342–350.
7. Budowle, B., J.A. Lindsey, J.A. De Cou, B.W. Koons, A.M. Giusti, and C.T. Comey. 1995. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQα using a multiplex amplification and typing procedure. J. Forensic Sci. 40:45-54.
8. Budowle, B., T.R. Moretti, S.J. Niezgoda, and B.L. Brown. 1998. CODIS and PCR-based short tandem repeat loci: law enforcement tools, p. 73–88. In Second European Symposium on Human Identification 1998, Promega Corporation, Madison, WI.
9. Butler, John M...”DNA Biology Review."In Forensic DNA typing biology, technology, and genetics of STR markers. 2nd ed. Amsterdam: Elsevier Academic Press, 2005. 36-50. Great overview of DNA Biology with a broad range of topics and some depth.
10. Butler, John M., Michael D. Coble, and Peter M. Vallone. 2007. "STRs vs. SNPs: Thoughts on the Future of Forensic DNA Testing." Forensic

Science, Medicine, and Pathology 3, no. 3: 200-205. Good for background on SNPS.

11. "Chelex® 100 Extraction Process."National Forensic Science TechnologyCenter.http://www.nfstc.org/pdi/Subject03/pdi_s03_m0301.htm. A guide to Chelex's use.
12. Ch. Gehrig, M. Hochmeister, B. Budowle, 2000. Swiss allele frequencies and haplotypes of 7 Y-Specific STRs, J. Forens. Sci. 45, 2 436-439.
13. Collins, P.J., L.K. Hennessy, C.S. Leibelt, R.K. Roby, D.J. Reeder, and P.A. Foxall. 2004. Developmental validation of a single-tube amplification of the 13 CODIS loci, D2S1338, D19S433, and amelogenin: the AmpFlSTR® Identifiler® PCR amplification kit. J. Forensic Sci. 49:1265 1277.
14. Comey, C.T. and B. Budowle. 1991. Validation studies on the analysis of the HLA-DQ alpha locus using the polymerase chain reaction. J. Forensic Sci. 36:1633-1648.
15. Clark, David P. 2005. "The Polymerase Chain Reaction." In Molecular Biology Understanding the Genetic Revolution. Burlington: Elsevier. 634-654. PCR in detail.
16. DNA Forensics: From RFLP to PCR-STR and Beyond;
17. http://www.forensicmag.com/articles/2004/09/dna-forensics-rflp-pcr-str-and-beyond.
18. "DNA Technology - The science behind DNA testing. "DNA Junction - Your resource for DNA testing information. http://www.dnajunction.com/technology. Source for DNA testing uses, but not for background information.
19. Edwards, A., A. Civitello, H.A. Hammond, and C.T. Caskey. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am. J. Hum.Genet. 49:746-756.

20. E.W. Hill, M.A. Jobling, D.G. Bradley. 2000. Y-chromosome variation and Irish origins, Nature 404. 351.
21. Giusti, A.M. and B. Budowle. 1995. Chemiluminescence-based detection system for human DNA quantitation and restriction fragment length polymorphism (RFLP) analysis. Appl. Theor. Electrophor. 5:89-98.
22. Gyllensten, U.B. and H.A. Erlich. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQ alpha locus. Proc. Natl. Acad. Sci. USA 85:7652-7656.
23. Jeffreys, A.J., V. Wilson, and S.L. Thein. 1985. Hypervariable minisatellite regions in human DNA. Nature 314:67-73.
24. Jeffereys, A.J. 1987. Hypervariable "minisatellite" regions in human DNA. Nature vol. 314, pp. 67-73.
25. Kayser M and de Knijff P. 2011.Improving human forensics through advances in genetics, genomics and molecular biology.Nat Rev Genet. 12: 179–192.
26. Krenke, B.E., A. Tereba, S.J. Anderson, E. Buel, S. Culhane, C.J. Finis, C.S. Tomsey, J.M. Zachetti, et al. 2002. Validation of a 16-locus fluorescent multiplex system. J. Forensic Sci. 47:773-785.
27. L. Roewer et al. 2000. Caucasian Y-STR haplotype reference database for forensic application, in G.F. Sensabaugh, P.J. Lincoln, B. Olaisen (eds.), Progress in Forensic Genetics 8 613-618.
28. L. Roewer et al. 2000. A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. Forensic Sci. Int. 114.31-43.
29. M.A. Jobling, A. Pandya, C. Tyler-Smith. 1997. The Y chromosome in forensic analysis and paternity testing, Int. J. Legal Med. 110. 118-124.

30. M. Kayser et al. 1997. Evaluation of Y chromosomal STRs: a multicenter study, Int. J. leg. Med. 110.125-133.
31. M. Kayser et al. 2000. Characteristics and frequency of germline mutations at microsatellites from the human Y chromosome revealed by direct observation in father/son pairs. Amer. J. Hum. Genet. 66. 1580-1588.
32. M. Prinz, K. Boll, H. Baum, B. Shaler. 1997. Multiplexing of Y chromosome specific STRs and performance of mixed samples, Forensic Sci. Int. 85. 209-218.
33. Mullis, K.B. and F.A. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. Methods in Enzymology vol. 155.335-350.
34. Murch, R.S. and B. Budowle. 1986. Applications of isoelectric focusing in forensic serology. J. Forensic Sci. 31:869-880.
35. Nakamura, Y., M. Leppert, P. O'Connell, R. Wolff, T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin and R. White. 1987. Variable number tandem repeat (VNTR) markers for human gene mapping. Science vol. 237.1616-1622.
36. P. de Knijff et al. 1997. Chromosome Y microsatellites: population genetic and evolutionary aspects, Int. J. legal Med. 110. 134-140.
37. P. E. Roffey, C. I. Eckhoff, J. L. Kuhl. 2000. A rare mutation in the amelogenin gene and its potential investigative ramnifications, J Forensic Sci 45(5). 1016-1019.
38. P.M. Schneider, S. Meuser, W. Waiyawuth, Ch. Rittner. 1998. Tandem repeat structure of the duplicated Y-chromosomal STR locus 385 and frequency studies in the German and three Asian populations, Forensic Sci. Int. 97. 61-70.
39. R. Decorte et al. 2000. STR (autosomal and Y-chromosome analysis reveals geographic differences in the Turkish population. In: G.F.

Sensabaugh, P.J. Lincoln, B. Olaisen (eds) Progress in Forensic Genetics 8. 215-217.

40. S.A. Miller, D.D. Dykes, H.F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells, Nucleic Acids Res. 16. 1215.
41. Saiki, R.K., S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction analysis for diagnosis of sickle cell anemia. Science 230:1350-1354.
42. Sensabaugh, G.F. 1995. Biochemical markers of individuality p. 338-415. In R. Saferstein (Ed.), Criminalistics—An Introduction to Forensic Science, 5th ed. Prentice Hall Education, Upper Saddle River, NJ.
43. Sequencing, Forensic Analysis and Genetic Analysis; http://www.atdbio.com/content/20/Sequencing-forensic-analysis-and-genetic-analysis.
44. Walsh, P.S., N. Fildes, A.S. Louie, and R. Higuchi. 1991. Report of the blind trial of the Cetus AmpliType HLA DQα forensic deoxyribonucleic acid (DNA) amplification and typing kit. J. Forensic Sci. 36:1551-1556.
45. Warshauer DH, Lin D, Hari K, Jain R, Davis C, et al. 2013.STRait Razor: A length-based forensic STR allele-calling tool for use with second generation sequencing data. Forensic SciInt Genet. 7: 409–417.
46. Whitaker, J.P., T.M. Clayton, A.J. Urquhart, E.S. Millican, T.J. Downes, C.P. Kimpton, and P. Gill. 1995. Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples. BioTechniques 18:670-677.
47. Wong, Z., V. Wilson, A.J. Jeffereys and S.L. Thein. 1986. Cloning a selected fragment from a human DNA "fingerprint": isolation of an

extremely polymorphic minisatellite. Nucleic Acids Research vol. 14, pp. 4605-4616.

48. Wyman, A.R. and R. White. 1980. A highly polymorphic locus in human DNA. Proc. Natl. Acad. Sci. USA 77:6754-6758.